

XX This protein produced by recombinant expression of genomic DNA isolated
CC from human prostate cancer PC-3 cells interacts with the response
CC elements and network of receptors in the thyroid hormone/retinoic acid
CC receptor subfamily. The products can be used in detection, diagnosis and
CC in screening assays for substances which interact with UNRPs for use in
CC diagnosis, drug design and therapeutic applications
XX Sequence 460 AA;

Alignment Scores:
Pred. No.: 1,27e-92 Length: 460
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAR74738 (1-460)

QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCAAGTCACAGTCGCGAGTCACCTGTGGGCGG 60
Db 168 LyeileAarglysgInGInGInGInSerGlnSerGlnSerProValGlyPro 187
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGCCCTGGGGCTTCCCTCGTGGATCTGAGGCGGC 120
Db 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 227
QY 181 CAGCAGTTGTGGGGCCCACTCAGTCAGTGCACAAACAGCTCTTCTCCGACCGCCCAAA 240
Db 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247
QY 241 GTACGGCTTGGCCCTGGGGCGGAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267
QY 301 GCCCACTTACGAGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
QY 361 GTCCTGTGTTTCTCGAGCTGGGGCGGAGGAGCAGATCGCCCTCTGAAAGGCATCCACT 420
Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLysAlaSerThr 307
QY 421 ATCGAGATCATGCTGTAGACACAGCCGCGCTACAAACAGCAGACAGAGTGTATCACC 480
Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327
QY 481 TTCTTGAG-GACTTTCACCTACAGCAAGGACGACTTCCACCGTGCAGGCTGCAGTGGAG 539
Db 328 PheLeuLysAspPheThrThrThrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTCATCAACCCCATCTCTCGATTTCTCGCGGCGCATCGCGGGCTTGGCCCTGGACGACGCT 599
Db 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 367
QY 600 GAGTACGCCCTGTCTATCGGCATCAACATCTTCTCGCGCGCACCGGCCCAACGTCGAGGAG 659
Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGGCCGCTGGAGGGCTGTGCAGCAGCCCTACCTGGAGGGCTGTCTCTACACCGCGC 719
Db 388 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 407
QY 720 ATCAAGAGCGCGCAGGACCACTCGCC 746
Db 408 IleLysArgProGlnAspGlnLeuArg 416

RESULT 2

AAR52980
ID AARS2980 standard; protein; 461 AA.
XX
AC AAR52980;
XX
DT 25-MAR-2003 (revised)
DT 06-OCT-1994 (first entry)
XX
DE Human recombinant steroid hormone receptor NERI protein.
XX
KW Steroid hormone receptor; NERI.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Binding-site 87..154
FT /note= "DNA binding area"
XX
PN WO9407916-A1.
XX
PD 14-APR-1994.
XX
PF 27-SEP-1993; 93WO-US009165.
XX
PR 07-OCT-1992; 92US-00958137.
XX
PA (NERI) MERCK & CO INC.
XX
PI Schmidt A, Rodan GA, Rutledge SJ, Vogel RL;
XX
DR WPI; 1994-135511/16.
DR N-PSDB; AA063134.
XX
PT New human steroid receptor NERI and corresp. DNA - used to develop assays
PT to identify steroid hormone agonists and antagonists.
XX
PS Claim 8; Page 31; 45pp; English.
XX
CC The NERI protein may be expressed in a CHO cell culture and used to
CC develop assays to identify steroid hormone agonists and antagonists.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 461 AA;

Alignment Scores:
Pred. No.: 1,27e-92 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAR52980 (1-461)

QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCAAGTCACAGTCGCGAGTCACCTGTGGGCGG 60
Db 169 LysileAarglysgInGInGInGInSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGGCTTCCCTCGTGGATCTGAGGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 228
QY 181 CAGCAGTTGTGGGGCCCACTCAGTCAGTGCACAAACAGCTCTTCTCCGACCGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCGCTTGGCCCTTGGGCGGAGACCCCGAGTCCCGAGATCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268

QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 DB 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
 QY 361 GTCCCTGGTTCTCGAGCTGGCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
 DB 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGACAGCCAGCGCTACAACCAGACAGAGATGTATCACC 480
 DB 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-CACCTTACCTACAGCAAGAGCGACTTCCACCGTGCAGGCGCTGCAGGTGGAG 539
 DB 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTATCAACCCCATCTTCAGATTCTCGCGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
 DB 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCCCTGCTATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
 DB 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGCGCGGCTGGAGCGGTTGACAGAGCCCTACGTGGAGCGGCTGCTGCTACACGCGC 719
 DB 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGCAGACGAGCTGGCGC 746
 DB 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 3

AAR97982
 ID AAR97982 standard; protein; 461 AA.

AC AAR97982;
 DT 25-SEP-1996 (first entry)
 XX Human steroid receptor NER.
 DE Steroid receptor; NER; TOFA; NGF potentiator; muscarinic agonist;
 KW dopamine D1 antagonist.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH 87..154
 FT Domain /note= "putative double zinc finger structure"

XX WO9613519-A1.

XX 09-MAY-1996.

XX 24-OCT-1995; 95WO-US013924.

XX 27-OCT-1994; 94US-00330283.

XX (MERI) MERCK & CO INC.
 PA (MEDI-) MEDICAL COLLEGE PENNSYLVANIA.
 XX

PI Friedman E, Holloway MK, Rodan GA, Rutledge SJ, Schmidt A;
 PI Vogel RU;

XX WPI; 1996-239449/24.

DR N-PSDB; AAT18996.

XX New isolated human steroid receptor NER - used to identify cpds. which
 PT can act, e.g., as potentiators of NGF, muscarinic agonists or dopamine D1
 PT antagonists.

PS Claim 5; Page 54-57; 72pp; English.

XX Steroid receptor NER (AAR97982) is a novel member of the nuclear receptor
 CC family. Its amino acid sequence was deduced from a gene (AAT18996)
 CC isolated from a human osteosarcoma SAOS-2/B10 cell library. A COS stable
 CC expression system allows prodn. of recombinant NER. NER is useful for
 CC identifying cpds. which act e.g. as potentiators of NGF, muscarinic
 CC agonists or dopamine D1 antagonists. TOFA (5-tetradecyloxy)-2-furan
 CC carboxylic acid) was found to activate NER and to be a potent potentiator
 CC of other drugs

XX Sequence 461 AA;

Alignment Scores:
 Pred. No.: 1-27e-92 Length: 461
 Score: 1231.00 Matches: 248
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 99.60% Mismatches: 1
 Query Match: 89.59% Indels: 1
 DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAR97982 (1-461)

QY 1 AAGATTCCGAAAACAGCAGCAGGAGTCAAGTCAAGTCCAGTCCAGTTCGCTGGGCGC 60
 DB 169 LysIleArgLysGlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 188
 QY 61 CAGGCGCAGCAGCAGCTCAGCCCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGAGGC 120
 DB 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
 QY 121 AGCCAGGCGTCCGGGAAGCGAGGCTGTCAGCTAACACAGCGCTCAAGAACTAATGATC 180
 DB 209 SerGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
 QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACAGCTCTCTTCGACACGCCCAAA 240
 DB 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
 QY 241 GTCAGCCCTGGCCCTTGGCGCAGACCCCGAGTCCCGAGATGCCCGCCAGCACCGCTTT 300
 DB 249 ValThrProTyrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
 QY 301 GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 DB 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
 QY 361 GTCCCTGGTTCTCGAGCTGGCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
 DB 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCGCTACAACCAGACAGACAGATGTATCACC 480
 DB 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-GACTTCACTTACAGCAAGGACGACTTCCACCGTGCAGGCGCTGCAGGTGGAG 539
 DB 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTATCAACCCCATCTTCAGATTCTCGCGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
 DB 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCCCTGCTATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
 DB 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGCGCGGCTGGAGCGGTTGACAGAGCCCTACGTGGAGCGGCTGCTGCTACACGCGC 719
 DB 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGCAGACGAGCTGGCGC 746

```

Db      409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 4
AAR98140
ID      AAR98140 standard; protein; 461 AA.
XX
XX
AC      AAR98140;
XX
XX      19-AUG-1996 (first entry)
XX
XX      NER receptor potentiator.
XX
XX      NER receptor; potentiator; steroid hormone receptor;
KW      G-protein coupled receptor; nerve growth factor; Alzheimer disease;
KW      ocular hypertension; schizophrenia; distonia; tardive dyskinesia;
KW      Gilles de la Tourette syndrome.
XX
XX      Homo sapiens.
XX
XX      W09613257-A1.
XX
XX      09-MAY-1996.
XX
XX      24-OCT-1995; 95WO-US013931.
XX
XX      27-OCT-1994; 94US-00330518.
XX
XX      (MERI ) MERCK & CO INC.
XX      (MEDI-) MEDICAL COLLEGE PENNSYLVANIA.
XX
XX      Friedman E, Holloway MK, Rodan GA, Schmidt A, Vogel RL;
XX
XX      WPI; 1996-239256/24.
XX      N-PSDB; AAT30031.
XX
XX      Use of steroid hormone receptor NER activators - for potentiating
XX      activity of modulator of G-protein coupled receptor.
XX
XX      Disclosure; Page 50-52; 63pp; English.
XX
XX      A novel human steroid hormone receptor (AAR98140), designated NER, is
XX      useful as a potentiator of ligands for other receptors, partic. G-protein
XX      coupled receptors. It was identified as the product of a DNA clone
XX      (AAR30031) isolated from an osteosarcoma SAOS-2/B10 library. Recombinant
XX      NER can be obtd. using a COS expression system. Cpds. that activate the
XX      NER receptor, such as TOFA (5-(tetradecyloxy)-2-furan carboxylic acid),
XX      potentiate the effects of nerve growth factor and may be useful in
XX      treatment of Alzheimer disease. They may also potentiate the effects of
XX      muscarinic agonists in the treatment of ocular hypertension, or dopamine
XX      D1 dopamine D1 antagonists in the treatment of psychoses and movement
XX      disorders
XX
XX      Sequence 461 AA;
XX
XX      Alignment Scores:
XX      Pred. No.: 1.27e-92 Length: 461
XX      Score: 1231.00 Matches: 248
XX      Percent Similarity: 99.60% Conservative: 0
XX      Best Local Similarity: 99.60% Mismatches: 1
XX      Query Match: 89.59% Indels: 1
XX      DB: 2 Gaps: 0
XX
XX      US-10-712-629B-2 (1-746) x AAR98140 (1-461)
XX
XX      1 AAGATTTCGGAACAGCAGCAGGAGTCAAGTCACAGTCGACAGTCACCTGTGGGCGG 60
XX
XX      169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
XX
XX      61 CAGGGCAGCAGCAGCTCAGCTCTGGGCGCTTCCCTGGTGATCTGAGGAGGC 120
XX
XX      189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
XX
XX      121 AGCCAGGGCTCCGGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGTC 180

```

```

Db      209 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
XX
XX      181 CAGCAGTTGGTGGCGGCCCACTGCAGTGCACAAACGCTCTCTTCCGACGACCCAAA 240
XX
XX      229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
XX
XX      241 GTCACGCCCTGGCCCTGGCGCCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
XX
XX      249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
XX
XX      301 GCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
XX
XX      269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
XX
XX      361 GTGCCCTGTGTTCTCTGAGCTGGCGGAGGAGACCATCGCCCTCTCTGAAGGCATCCACT 420
XX
XX      289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
XX
XX      421 ATCGAGATCATGCTGCTAGAGACAGCCAGCGCTTCAACACGAGACAGACAGTGTATCACC 480
XX
XX      309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
XX
XX      481 TTCTTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCGAGCGCTGCAGGTGGAG 539
XX
XX      329 PheLeuLysAspPheThrTyrSerLysAspPhePheHisArgAlaGlyLeuGlnValGlu 348
XX
XX      540 TTCATCAACCCCATCTTTCGAGTTCTCGCGGCGCATCGCGCGCTGGCGCTGGACGACGT 599
XX
XX      349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLysGlyLeuAspAspAla 368
XX
XX      600 GAGTAGCGCCCTGCTCATGCCATCAACATCTTCTCGCGCCGACGCGGCCCAACGTCGAGGAG 659
XX
XX      369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
XX
XX      660 CGGGCGCGCTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGCTGTCTCTACACGCGC 719
XX
XX      389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
XX
XX      720 ATCAAGAGCGCGCAGGACGACGCTGCGC 746
XX
XX      409 IleLysArgProGlnAspGlnLeuArg 417
XX
XX      RESULT 5
XX      ABG31015
XX      ID      ABG31015 standard; protein; 461 AA.
XX
XX      AC      ABG31015;
XX
XX      21-OCT-2002 (first entry)
XX
XX      Human LX receptor beta (LXRB).
XX
XX      LX receptor beta; LXRB; targeting construct; transgenic animal; diabetes;
XX      impaired glucose tolerance; reduced blood insulin level; hypoactivity;
XX      lethargy; human.
XX
XX      Homo sapiens.
XX
XX      W0200257438-A2.
XX
XX      25-JUL-2002.
XX
XX      11-DEC-2001; 2001WO-US047989.
XX
XX      11-DEC-2000; 2000US-0254801P.
XX      31-JUL-2001; 2001US-0309404P.
XX      10-DEC-2001; 2001US-00013823.
XX
XX      (DELT-) DELTAGEN INC.
XX
XX      Guenther C, Phillips R, Allen KD, Zhang Q, Baribault H;

```



```
DB: 8 Gaps: 0
US-10-712-629B-2 (1-746) x ADU14182 (1-461)
QY 1 AAGATTCGGAACAGCAGCAGGAGTACAGTCACAGTCGCGAGTCACCTGTGGGCGG 60
DB 169 LyeileArglysgInGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTCGTGGATCTGAGGCAGGC 120
DB 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCACGCTAACAGCGCTCAAGAACTAATGATC 180
DB 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 228
QY 181 CAGCAGTTGGTGGCGCCCACTGCAGTGCAACAAAGCTCTCTCCGACGAGCCCA 240
DB 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGCCCTGGCGCAGACCCAGTCGCCAGATGCCGCGCAGCAACGCTTT 300
DB 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCAGCTTACGAGCTGGCCATCTCAGTCCAGGAGATCTGGACTTCGCTAAGCAA 360
DB 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValaAspPheAlaLysGln 288
QY 361 GTCCTGGTTCTCGAGCTGGCGGAGGACGAGTCCGCTCTCGAAGGCATCCACT 420
DB 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGACAGCAGCGCTTACACACAGCAGACAGAGTGTATCACC 480
DB 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTTCTGAG-CAGTTACCTACAGCAAGGACGACTTCCACCGTCGAGCGCTTCAGGTGGAG 539
DB 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTTATCAACCCATCTCGAGTCTCGCGGCAATCGCGGCTGGCGCTGGAGCGCT 599
DB 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGTCTCATCGCATCAACATCTTCTCGCGCGACCGCGCCAAACGTCAGGAG 659
DB 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGCGCGGTGGAGCGTTGACAGCCCTTACGTGAGGCGCTGTGCTCTACACCGCG 719
DB 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 408
QY 720 ATCAGAGCGCGGAGGACGAGTCCGC 746
DB 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 7
ADP05729
ID ADP05729 standard; protein; 461 AA.
XX
AC ADP05729;
XX
DB 26-AUG-2004 (first entry)
XX
DE Human nuclear receptor protein SeqID103.
XX
KW disease risk; disorder risk; mutation; polymorphism;
KW nuclear receptor protein; antibacterial; antithyroid; cardiovascular-Gen;
KW cytosolic; dermatological; eating-disorders-Gen; gastrointestinal-Gen;
KW gynaecological; hepatotropic; immunosuppressive; muscular-Gen;
KW nephrotropic; osteopathic; virucide; adrenal gland; colon;
KW cardiovascular; intestine; kidney; liver; lung; muscular; ovary; blood;
KW prostate; skin; spleen; stomach; testes; thymus; thyroid; uterus;
```

KW pancreas; bone; joint; breast; immune system; metabolic; nutritive disease; human.

OS Homo sapiens.

XX WO2004045359-A2.

XX 03-JUN-2004.

XX 12-NOV-2003; 2003WO-US036229.

XX 14-NOV-2002; 2002US-0426305P.

XX (NURA-) NURA INC.

XX Gaitanaris GA, Bergmann JE, Gracarov A, Hohmann J, Li F;

PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;

XX WPI; 2004-449627/42.

XX N-PSDB; ADP05730.

XX Determining an increased risk for e.g. colon, brain or breast disease or disorder, by detecting a mutation or polymorphism in the nuclear receptor gene, or measuring expression or biological activity level of the nuclear receptor.

XX Claim 1; SEQ ID NO 103; 508pp; English.

XX This invention relates to a novel method of determining whether a patient has an increased risk for developing a disease or disorder which comprises determining the presence of a mutation or polymorphism in the patient's gene encoding a nuclear receptor protein or measuring the expression or level of biological activity of a nuclear receptor polypeptide in the patient or in a cell of the patient. The invention may be useful for the development of compounds with an antibacterial, antithyroid, cardiovascular-Gen, cytostatic, dermatological, eating-disorders-Gen, gastrointestinal-Gen, gynaecological, hepatotropic, immunosuppressive, muscular-Gen, nephrotropic, osteopathic or virucide activity. The method is useful for determining whether a patient has an increased risk for developing a disease or disorder. The nucleic acid encoding a nuclear receptor polypeptide, an expression vector comprising the nucleic acid operably linked to a promoter, or a compound that modulates the biological activity of a nuclear receptor polypeptide, is useful for treating or preventing a disease or disorder of the adrenal gland, colon, cardiovascular, intestine, kidney, liver, lung, muscular, ovary, blood, prostate, skin, spleen, stomach, testes, thymus, thyroid, uterus, pancreas, bone and joints, breast, or immune system, or metabolic or nutritive disease or disorder. The present sequence is that of a nuclear receptor protein which may be used in the method of the invention.

XX Sequence 461 AA;

Alignment Scores:
Pred. No.: 1.27e-92 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 8 Gaps: 0

US-10-712-629B-2 (1-746) x ADP05729 (1-461)

QY 1 AAGATTCGGAACAGCAGCAGGAGTACAGTCACAGTCGCGAGTCACCTGTGGGCGG 60

DB 169 LyeileArglysgInGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188

QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTCGTGGATCTGAGGCAGGC 120

DB 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208

QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCACGCTAACAGCGCTCAAGAACTAATGATC 180

Db 209 SerGlnGlySerGlyGluGlyGluValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCACTGCAGTCAACAAACAGCTCTCTCCGACGACCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGCCCTGGCGGAGACCCCACTGCCAGATCCCGCCACGACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTACAGGAGCTGGCCATCATCTCAGTCCAGGAGATCTGGACTTCGTAGCAA 360
Db 269 AlaHisPheThrGlnLeuAlaIleLeuSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCTCTGTTCTCTGACGTGGCGGAGGACACAGATCCCTCTCTGAGGACATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCTACAACACGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACGTCAGCGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCCATCGCGGCTGGCGCTGGACGACGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgGluGlyLeuAspAla 368
QY 600 GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGCGCCGACCGCCCAACGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGCGCGCGCTGAGCGCTGACGACGCGCTAGCTGAGCGCTGTCTGTCTACAGCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCGCGCAGGACGAGTGGCG 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 8

ADQ48625
ID ADQ48625 standard; protein; 461 AA.
XX
AC ADQ48625;
XX
DT 07-OCT-2004 (first entry)
XX
DE Wild type human liver X receptor beta protein.
XX
KW liver X receptor beta; LXRbeta; ligand binding domain;
KW protein crystallography; X-ray crystallography; drug screening assay;
KW human; protein co-ordinate data.
XX
OS Homo sapiens..
XX
FH Key Location/Qualifiers
FT Region 220..461
FT /note= "region is specifically claimed"
XX
PN WO2004058819-A2.
XX
PD 15-JUL-2004.
XX
PF 24-DEC-2003; 2003WO-IB006412.
XX
PR 24-DEC-2002; 2002GB-00030177.
XX
PA (KARO-) KARO BIO AB.
XX
PI Farnegardh M, Bonn T, Sun S, Ljunggren J, Ahola H, Carlquist M;

XX WPI; 2004-543437/52.
DR
XX
PT New crystal comprising at least 150 amino acid residues of the LXRbeta
PT ligand-binding domain, useful in X-ray crystallography, and for drug
PT screening assay.
XX
PS Claim 31; SEQ ID NO 1; 378pp; English.
XX
CC The invention relates to a crystal comprising at least 150 amino acid
CC residues of the liver X receptor beta (LXRbeta) ligand binding domain.
CC The ligand binding domain preferably comprises amino acid residues
CC Ser242, Phe268, Thr271, Thr272, Leu274, Ala275, Ser278, Ile309, Met312,
CC Leu313, Glu315, Thr316, Arg319, Ile327, Leu329, Leu330, Tyr335, Phe340,
CC Leu345, Phe349, Ile350, Ile353, Phe354, His435, Gln438, Val439, Leu442,
CC Leu449, Leu453, Trp457, according to the co-ordinate tables or homologue
CC of the molecule or molecular complex, where the homologue has a root mean
CC square deviation from the backbone atoms of the amino acids of not more
CC than 1.5A. Also disclosed is an isolated protein consisting essentially
CC of the amino acid 220-461 of the 461 LXRbeta protein. The crystal is
CC useful in X-ray crystallography, and for drug screening assays. This
CC sequence corresponds to the full length human LXRbeta protein.
XX
SQ Sequence 461 AA;
Alignment Scores:
Pred. No.: 1-27e-92 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: Gaps: 0
US-10-712-629B-2 (1-746) x ADQ48625 (1-461)
QY 1 AAGATTTCGAAACAGCAGCAGGAGTCAAGTCAAGTCCAGTCCAGTCTGTTGGGCGC 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGCGCTGGGGCTTCCCTGGTGGATCTGAGCGAGC 120
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGGAAGCGAGGGTGTCCAGTCAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACAGCTCTCTTCCGACGACCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGCCCTGGCGGAGACCCCACTGCCAGTCCGAGATGCCCGCCAGCAGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTACAGGAGCTGGCCATCATCTCAGTCCAGGAGATCTGGACTTCGTAGCAA 360
Db 269 AlaHisPheThrGlnLeuAlaIleLeuSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCTCTGTTCTCTGACGTGGCGGAGGACGACTTCCACGTCAGCGCTGTCTGTAGGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCTACAACACGACGAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACGTCAGCGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGCCATCGCGGCTGGCGCTGGACGACGCT 599

Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCGCTGCTCATCGCATCAACATCTTCTCGCGCGACCGGCCCAACGTCGAGGAG 659
 Db 369 GluTyrAlaLeuLeuLeuAlaIleAsnIlePheSerAlaAspArgProAsnValGlu 388
 QY 660 CCGGCGCGCTGAGCGCTTGCAGCAGCCCTACGTGAGCGCTGCTGCTTACAGCGC 719
 Db 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGAGCAGCAGCTGCGC 746
 Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 9

AEB94510

ID AEB94510 standard; protein; 461 AA.

XX

AC AEB94510;

XX

DT 06-OCT-2005 (first entry)

XX

DE Human LXR protein - SEQ ID 12.

XX

XX lipid metabolism; insulin resistance;
 KW genetically engineered microorganism; screening; LXR.
 XX

OS Homo sapiens.

XX

PN JP2005204550-A.

XX

PD 04-AUG-2005.

XX

XX 21-JAN-2004; 2004JP-00013728.

XX

XX 21-JAN-2004; 2004JP-00013728.

PR

XX (YAMA) YAMANOUCHI PHARM CO LTD.

PA

XX Ida M, Ogino A, Endo H, Nishijima S, Inoki Y;

PI

XX WPI; 2005-546101/56.

XX

DR N-PSDB; AEB94509.

XX

XX Screening lipid metabolism or insulin resistance improvement drug, by
 PT contacting test substance with cell capable of controlling ABCA1
 PT expression, analyzing change in liver X receptor dependent
 PT transcriptional activity.

XX

PS Disclosure; SEQ ID NO 12; 25pp; Japanese.

XX

XX The invention comprises a method of screening a drug for improving lipid
 CC metabolism or insulin resistance. The method involves contacting a test
 CC substance with a cell transformed by an expression vector encoding an LXR
 CC transcription coupling factor capable of controlling ABCA1 expression,
 CC and analyzing the effect of test substance on change in liver X receptor
 CC dependent transcriptional activity. The method of the invention is useful
 CC for screening a drug for improving lipid metabolism or insulin
 CC resistance. The present amino acid sequence represents a human LXR
 CC protein.

XX

SQ Sequence 461 AA;

Alignment Scores:

Pred. No.:	1-27e-92	Length:	461
Score:	1231.00	Matches:	248
Percent Similarity:	99.60%	Conservative:	0
Best Local Similarity:	99.60%	Mismatches:	1
Query Match:	89.59%	Indels:	1
DB:	9	Gaps:	0

US-10-712-629B-2 (1-746) x AEB94510 (1-461)

QY 1 AAGATTGGAACACAGCAGCAGGAGTCAAGTCAAGTCCAGTCACTGTGGGGCCG 60
 Db 169 LysIleArgLysGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
 QY 61 CAGGCGCAGCAGCAGCTCAGCCCTCGGCTGGGCTTCCCTGGTGGATCTGAGGCGAGC 120
 Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
 QY 121 AGCCAGGGTCCGGGGAAGCGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
 Db 209 SerGlnGlySerGlyGlyGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
 QY 181 CAGCAGTTGGTGGCGGCCCAACTGACGTGCAACAAACGCTCTTCTCCACACGCCA 240
 Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
 QY 241 GTCAGCCCTGGCCCTGGCGCAGACCCCGAGATGCCCGCAGCAGCAACCTTT 300
 Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 268
 QY 301 GCCCACTTCACGAGCTGGCCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
 QY 361 GTGCTGTGTTCTCGAGCTGGCGGAGGAGCAGATCGCCCTCCTGAGGATCCACT 420
 Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
 QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCTACAACACAGCAGACAGAGTGTATCACC 480
 Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
 QY 481 TTCTTGAG-GACTTCACCTACAGCAGAGCAGCTTCCACCGTCAGCGCTGCAAGTGAG 539
 Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
 QY 540 TTCATCAACCCCATCTTCGAGTTCGCGGGCCATCGCGGCTGGGCTGGACGAGCT 599
 Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
 QY 600 GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTGCAGGAG 659
 Db 369 GluTyrAlaLeuLeuLeuAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
 QY 660 CCGGCGCGCTGAGCGCTTGCAGCAGCCCTACGTGAGCGCGCTGTGCTCTACAGCGC 719
 Db 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
 QY 720 ATCAAGAGCGCGCAGCAGCAGCTGCGC 746
 Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 10

ADL12926

ID ADL12926 standard; protein; 461 AA.

XX

XX AC ADL12926;

XX

DT 06-MAY-2004 (first entry)

XX

DE Human steroid-induced C3A liver cell protein #105.

XX

XX Hepatotropic; Gene therapy; Wilson disease; liver disorder;
 KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
 XX

XX Homo sapiens.

XX

XX OS US6673549-B1.

XX

XX PN 06-JAN-2004.

XX

XX 12-OCT-2001; 2001US-00976594.

XX

PR 12-OCT-2000; 2000US-0240409P.
XX (INCY-) INCYTE CORP.
PA Furness LM, Buchbinder JL;
PI WPI; 2004-068610/07.
XX
DR Combination useful for preparing a composition for treating liver
XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
PT
XX Disclosure; SEQ ID NO 655; 141bp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a cDNA. The sample is from a subject with
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 461 AA;

Alignment Scores:
Pred. No.: 2,7e-92 Length: 461
Score: 1227.00 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 2
Query Match: 89.30% Indels: 1
DB: Gaps: 0

US-10-712-629B-2 (1-746) x ADL12926 (1-461)

QY 1 AAGATTCCGAACAGCAGCAGCAGGAGTCACAGTCACAGTCACCTGCGGCGC 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188

QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGCGATCTGAGGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208

QY 121 AGCCAGGGCTCCGGGGAAGCCAGGCTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyGlyValGlnGlnLeuThrAlaAlaGlnGlnLeuMetIle 228

QY 181 CAGCAGTTGGTGGCGGCCCAACTCAGTGTCAACAAACGCTCTCTCCGACCGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248

QY 241 GTCAGCGCTCCGCGCTGGCGCAGACCCCGAGTCCGAGATCCCGCCAGCAAGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268

QY 301 GCCCACTTCACCGAGCTGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGlnLeuValAspPheAlaLysGln 288

QY 361 GTGCTGTGTTCTCCTCAGCTGGCGGGAGGACACAGATCGCCCTCTCGAAGCATCCACT 420
XX

Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAaspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGAGACAGCAGCGCTACAACCCACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGlnThrGluCysIleThr 328
QY 481 TTCTTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTGCAGGCGCTGCAGGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTATCAACACCCATCTTCAGTTCTCGCGGCGCATCGCGGCTGGGCTGACGAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCGTGGAGCGCTTGAGCAGCAGCCCTACGTGGAGCGCTGTCTCTACAGCGCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluThrLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCGCGCAGGACCGAGCTGCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 11
AAR96234
ID AAR96234 standard; protein; 461 AA.
AC AAR96234;
XX
DT 06-NOV-1996 (first entry)
XX
DE Human foetal lung steroid hormone receptor analogue ECDN.
XX
KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW ECDN protein; cancer; screening; binding molecule; recombinant;
KW identification; anticancer drug; cancerous tissue; primer; probe;
KW antibody; immunohistochemical assay.
XX
OS Homo sapiens.
XX
PN WO9609324-A1.
XX
PD 28-MAR-1996.
XX
XX 21-SEP-1995; 95WO-JP001909.
XX
PR 21-SEP-1994; 94JP-00226270.
XX
PA (CANC-) CANCER INST.
PA (EISA) EISA CO LTD.
XX
PI Nakamura Y, Saito H;
XX
XX WPI: 1996-188403/19.
DR N-PSDB; AAT27616.
XX
PT ECDN protein, a steroid hormone receptor analogue from human foetal lung
PT - is expressed in cancer cells and is useful for cancer diagnosis and
PT drug development.
XX
PS Claim 1; Page 21-25; 43pp; Japanese.
XX
CC The present sequence is the human foetal lung derived steroid hormone
CC receptor analogue protein, ECDN. In various cancer cells a variant ECDN
CC protein, designated ECDN small mol. (ECDNsm) protein, is expressed.
CC Therefore screening for ECDN and ECDNsm protein binding mols., using
CC recombinant ECDN and ECDNsm proteins will be useful in the identification
CC of candidate anticancer drugs. Gene expression of ECDN and ECDNsm
CC proteins in normal and cancerous tissues can be studied using primers and
CC probes derived from ECDN and ECDNsm protein cDNA. Antibodies which

CC recognise ECDN and ECDNm proteins can be used in ECDN and ECDNm protein
CC immunohistochemical assays

XX Sequence 461 AA;

SQ Sequence 461 AA;

Alignment Scores:
Pred. No.: 3,27e-92 Length: 461
Score: 1226.00 Matches: 247
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.23% Indels: 1
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAR96234 (1-461)

QY 1 AGATTTCGGAACAGCAGCAGGAGTACAGTCACTCCAGTCACTGCTGGGGCGG 60
Db 169 Lys11eArgLysGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGGTGCATCTGAGGCGG 120
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCTCCGGGGAAGCGAGGCTGTCCAGTCAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCACTGCAGTGCACAAACAGCTCTTCCGACCGAGCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGCCCTGGGCGCAGACCCCGAGTCCCGAGATCCCGCGCAGCAACGCTTT 300
Db 249 ValThrProtrProleuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTGGTTCTCTGAGTGGCGGGAGGACAGATCGCCCTCTGAGGCGATCCACT 420
Db 289 ValProGlyPheLeuGlnGluGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCTACAACCCAGCAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTGCAGGCTGCGAGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTTCATCAACCCCATCTCGAGTTCTCGGGCCATCGGGCGCTGGGCTGGACGAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLysGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGTCTATCGGCATCAACATCTTCTCGGCCGACCGGCCCAACGTCGAGG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCGTGGAGCGCTTGCAGCAGCCCTTACGTGAGGCGCTGTGCTCTACACCGCG 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 408
QY 720 ATCAAGGCGCGGAGGACGAGCTG 743
Db 409 IleLysArgProGlnAspGlnLeu 416

RESULT 12

ID AAW25034

XX AAW25034 standard; protein: 460 AA.

AC

XX AAW25034;

XX

DT 25-MAR-2003 (revised)
DT 15-OCT-1997 (first entry)
XX
DE Human ubiquitous nuclear receptor polypeptide.
XX Ubiquitous nuclear receptor polypeptide; DNA binding domain; therapy;
KW drug design; diagnosis.
OS Homo sapiens.
XX US5639616-A.
PN 17-JUN-1997.
PD 18-NOV-1994; 94US-00342411.
PF 10-NOV-1993; 93US-00152003.
PR (ARCH-) ARCH DEV CORP.
PA Song C, Liao S;
PI WPI; 1997-332043/30.
XX N-PSDB; AAW25034.
DR DNA encoding ubiquitous nuclear receptor polypeptide(s) - for producing
recombinant polypeptide(s) and detecting RNA.
XX Claim 1; Col 71-74; 67pp; English.
XX AAW25034 shows the amino acid sequence of a human ubiquitous nuclear
receptor polypeptide (UR). The UR is useful in assays designed to select
substances which interact with the UR which may potentially be of use
therapeutically. UR is also used in diagnosis and drug design. (Updated
on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 460 AA;

Alignment Scores:

Pred. No.: 5.77e-92 Length: 460
Score: 1223.00 Matches: 246
Percent Similarity: 99.20% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 2
Query Match: 89.01% Indels: 0
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x AAW25034 (1-460)

QY 1 AGATTTCGGAACAGCAGCAGGAGTCACTCAGTCACTGCTGGGGCGG 60
Db 168 Lys11eArgLysGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 187
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGGTGCATCTGAGGCGG 120
Db 188 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207
QY 121 AGCCAGGCTTCGGGGAAGCGAGGCTGTCCAGTCAACAGCGCTCAAGAACTAATGATC 180
Db 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 227
QY 181 CAGCAGTTGGTGGCGGCCCACTGCAGTGCACAAACAGCTCTTCCGACCGAGCCAAA 240
Db 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247
QY 241 GTCACGCCCTGGGCCCTGGGCGCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db 248 ValThrProtrProleuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267
QY 301 GCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGAGCTTCGCTAAGCAA 360
Db 268 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 287
QY 361 GTGCTGGTTCTCTGAGCTGGGCGGAGGACGATCGCCCTCTCTGAAGGCATCCACT 420

Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
QY 421 ATCGAGATCATGCTAGAGACAGCCAGCGCTACAAACCAGACAGAGTGTATCACC 480
Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db 328 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 347
QY 540 TTATCAACCCCATCTTCGAGTTCTCGCGGCATCGCGGCTGGCGCTGGACGACGCT 599
Db 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 367
QY 600 GAGTACGCTGCTCATCGCATCAACATCTTCTCGCGCAGCCGCGCCACGTCGAGGAG 659
Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
QY 660 CCGGCGCGCTGGAGCGTTGACAGCAGCCCTACGTGGAGCGCTGTGTCTACACGCGC 719
Db 388 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
QY 720 ATCAAGAGCGCGAGACCAAGCTGGCG 746
Db 408 TyrLysArgProGlnAspGlnLeuArg 416

RESULT 13

AA96235
ID AA96235 standard; protein; 364 AA.
XX
AC AA96235;
XX
DT 06-NOV-1996 (first entry)
XX
DE Steroid hormone receptor analogue ECDN small mol. variant.
XX
KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW ECDN protein; cancer; screening; binding molecule; recombinant;
KW identification; anticancer drug; cancerous tissue; primer; probe;
KW antibody; immunohistochemical assay; variant; small molecule; ECDNm.
XX
OS Homo sapiens.
XX
PN WO9609324-A1.
XX
PD 28-MAR-1996.
XX
PF 21-SEP-1995; 95WO-JP001909.
XX
PR 21-SEP-1994; 94JP-00226270.
XX
PA (CANC-) CANCER INST.
PA (EISA) EISAI CO LTD.
XX
PI Nakamura Y, Saito H;
XX
DR WPI; 1996-188403/19.
DR N-PSDB; AAT27617.
XX

XX ECDN protein, a steroid hormone receptor analogue from human foetal lung
PT - is expressed in cancer cells and is useful for cancer diagnosis and
PT drug development.
XX
PS Claim 3; Page 25-28; 43pp; Japanese.
XX

CC The present sequence is the variant of the human foetal lung derived
CC steroid hormone receptor analogue protein ECDN, designated ECDN small
CC mol. (ECDNm) protein. ECDNm protein is expressed in various cancer
CC cells, therefore screening for ECDNm protein binding moles., using
CC recombinant ECDNm proteins will be useful in the identification of
CC candidate anticancer drugs. Gene expression of ECDNm proteins in
CC cancerous tissues can be studied using primers and probes derived from

CC ECDNm protein cDNA. Antibodies which recognise ECDNm proteins can be
CC used in ECDNm protein immunohistochemical assays
XX
SQ Sequence 364 AA;

Alignment Scores:

Pred. No.: 8.08e-92 Length: 364
Score: 1221.00 Matches: 246
Percent Similarity: 98.80% Conservative: 0
Best Local Similarity: 98.80% Mismatches: 3
Query Match: 88.86% Indels: 1
DB: Gaps: 0
2

US-10-712-629B-2 (1-746) x AA96235 (1-364)

QY 1 AAGATTCCGNAACAGCAGCAGGAGTCACTACAGTCGAGTCACCTGTGGGGCGG 60
Db 72 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 91
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTCCCTGTGGATCTGAGGACGGC 120
Db 92 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 111
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db 112 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 131
QY 181 CAGCAGTTGGTGGCGGCCCACTGCAGTGCAAAACGCTCTCTCCGACGACCCAAA 240
Db 132 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 151
QY 241 GTCACGCTCCGCGCTGGCGCAGACCCCGAGTCCGAGATCCCGCCAGCAACGCTTT 300
Db 152 ValThrProTyrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 171
QY 301 GCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 172 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 191
QY 361 GTGCTGTGTTCTCGAGTGGCGCGGAGGACCATGCGCTCTCTGAAGGATCCACT 420
Db 192 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleProLeuLeuAsnAlaSerThr 211
QY 421 ATCGAGATCATGCTAGACAGCAGCGCTACAAACCAGACAGAGTGTATCACC 480
Db 212 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 231
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db 232 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 251
QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGCATCGCGGCTGGCGCTGGAGCGCT 599
Db 252 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 271
QY 600 GAGTACGCTGCTCATCGCATCAACATCTTCTCGCGCAGCCGCGCCACGTCGAGGAG 659
Db 272 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 291
QY 660 CCGGCGCGCTGGAGCGTTGACAGCAGCCCTACGTGGAGCGCTGTGTCTACACGCGC 719
Db 292 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 311
QY 720 ATCAAGAGCGCGAGACCAAGCTGGCG 746
Db 312 IleLysArgProGlnAspGlnLeuArg 320

RESULT 14

AD85628
ID AD85628 standard; protein; 459 AA.
XX
AC AD85628;
XX

DT 05-MAY-2005 (first entry)
 XX Human LXRbeta protein.
 XX inflammatory bowel disease; LXRbeta; nuclear hormone receptor;
 KW metabolism; cholesterol; pharmaceutical; Crohn's disease;
 KW ulcerative colitis; colitis; ischemia; gastrointestinal-gen.;
 KW antiinflammatory; antiulcer; receptor.
 XX
 OS Homo sapiens.
 XX
 PN W02005013946-A2.
 XX
 PD 17-FEB-2005.
 XX
 PF 27-JUL-2004; 2004WO-EP008426.
 XX
 PR 28-JUL-2003; 2003US-0490614P.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Goto Y, Kikkawa H, Kinoshita M;
 XX
 DR WPI; 2005-172945/18.
 DR N-PSDB; ADX85627.
 XX
 PT Preventing and/or treating inflammatory bowel diseases, such as Crohn's
 PT disease, ulcerative colitis and inflammatory colitis, comprises
 PT administering LXR agonist.
 XX
 PS Disclosure; SEQ ID NO 4; 55pp; English.
 XX
 CC The invention relates to a method of treating or preventing inflammatory
 CC bowel disease (IBD) in a mammal. The method comprises administering a LXR
 CC agonist, or its pharmaceutical salt, solvent, or physiologically
 CC functional derivative. The LXR agonist is a compound of formula (I) or
 CC (II), as defined in the specification. The LXR agonist is defined as any
 CC compound that enhances the biological activity of LXRalpha and/or
 CC LXRbeta. LXRalpha and LXRbeta (collectively known as LXR) are nuclear
 CC hormone receptors that regulate the metabolism of lipids such as
 CC cholesterol. The LXR agonists are useful in a pharmaceutical composition
 CC for preventing and/or treating inflammatory bowel diseases, such as
 CC Crohn's disease, ulcerative colitis, and inflammatory colitis caused by
 CC bacteria, ischemia, radiation, drugs or chemical substances. This
 CC sequence represents human LXRbeta protein.
 XX
 SQ Sequence 459 AA;
 Alignment Scores:
 Pred. No.: 2,39e-91 Length: 459
 Score: 1215.50 Matches: 247
 Percent Similarity: 99.20% Conservative: 0
 Best Local Similarity: 99.20% Mismatches: 1
 Query Match: 88.46% Indels: 2
 DB: 9 Gaps: 1
 US-10-712-629B-2 (1-746) x ADX85628 (1-459)
 QY 1 AAGATTCGGAACAGCAGCAGCAGCAGTCACAGTCACAGTCGAGTCACCTGTGGGCGC 60
 Db 169 Lysilearglys---GlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 187
 QY 61 CAGGGCAGCAGCAGCTCAGCTTCGGGCGCTTCCCTGGTGGATCTGAGCAGGC 120
 Db 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGlyAlaGly 207
 QY 121 AGCAGGGCTCCGGGAGCGAGGTGTCAGCTACAGCGCTCAAGACTAATCATC 180
 Db 208 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMecile 227
 QY 181 CAGCAGTTGGTGGCGGCCAACCTCAGTCAGCAACAGCGCTCTTCTCCGACCGCCAAA 240
 Db 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247

QY 241 GTCAACGCTGCGCCCTGGGCGCAGACCCCCAGTCCCGAGATGCCCGCAGCAACGCTTT 300
 Db 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267
 QY 301 GCCCACTTCACGAGCTGGCCCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
 Db 268 AlaHisPheThrGlnLeuAlaIleIleSerValGlnGluIleValAlaPheAlaLysGln 287
 QY 361 GTGCTCTGTTTCTCTGAGCTGGGCGGAGGACCATCGCTCTCTGAAGGCATCCACT 420
 Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307
 QY 421 ATCAGATCATGCTGTAGACAGCCAGCGGTACACCAACGAGACAGAGTGTATCACC 480
 Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327
 QY 481 TTCTTCGAG-GACTTTCACCTACAGCAAGAGCAGCTTCCACCGTCAGCGCTGCAAGTGGAG 539
 Db 328 PheLeuLysAspPheThrTyrSerLysAspPhePheHisArgAlaGlyLeuGlnValGlu 347
 QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCGCATCGCGGCTGGGCTTGGACGAGCT 599
 Db 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 367
 QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGCAGCCGCAACGCTGCAAGG 659
 Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387
 QY 660 CCGGGCGCGGTGGAGCGGTTGACAGAGCCCTACGTGGAGCGCTGTGTCTCTACAGCGCG 719
 Db 388 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
 QY 720 ATCAAGAGCGCGCAGCAGCAGCTGCGC 746
 Db 408 IleLysArgProGlnAspGlnLeuArg 416
 RESULT 15
 ID ABR82544
 AC ABR82544;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human LXR beta polypeptide.
 KW Pyridone; antilipemic; antidiabetic; antiarteriosclerotic; litholytic;
 KW hepatotropic; antiseborrheic; dermatological; antiparkinsonian; cardiant;
 KW nootropic; neuroprotective; antiinflammatory; cytostatic; anorectic;
 KW cardiovascular; nephrotropic; antithyroid; ophthalmological; hypotensive;
 KW cerebroprotective; immunosuppressive; nuclear receptor; liver X receptor;
 XX LXR beta; human.
 OS Homo sapiens.
 XX
 PN W02003059884-A1.
 XX
 PD 24-JUL-2003.
 PF 20-DEC-2002; 2002WO-US041306.
 XX
 PR 21-DEC-2001; 2001US-0342707P.
 XX
 PA (XCEP-) X-CEPTOR THERAPEUTICS INC.
 XX
 PI Bayne CD, Johnson AT, Lu S, Mohan R, Griffith RC;
 XX
 DR WPI; 2003-646029/61.
 DR N-PSDB; ACF36088.
 XX
 PT New N-substituted pyridone compounds useful for the treatment of e.g.
 PT hypercholesterolemia, hyperlipoproteinemia, hyperglycemia, diabetes

mellitus, dyslipidemia or atherosclerosis.

Disclosure; Page 517-519; 544pp; English.

The invention relates to N-substituted pyridone compounds of specified formula. The compounds act as modulators of nuclear receptors, orphan nuclear receptors and liver X receptor (LXRalpha or LXRbeta). They are useful for the treatment, prevention or amelioration of symptoms of nuclear receptor mediated diseases or disorders which are modulated or affected by nuclear receptor activity or in which nuclear receptor activity is implicated e.g. hypercholesterolemia, hyperlipoproteinemia, hypertriglyceridemia, lipodystrophy, hyperglycemia, diabetes mellitus, dyslipidemia, atherosclerosis, gallstone disease, acne vulgaris, acneiform skin conditions, diabetes, Parkinson's disease, Alzheimer's disease, cancer, inflammation, immunological disorders, lipid disorders, obesity, conditions characterized by a perturbed epidermal barrier function, conditions of disturbed differentiation or excess proliferation of the epidermis or mucous membrane, cardiovascular disorders, symptoms of a disease or disorder affected by cholesterol, triglyceride or bile acid levels, symptoms of hypocholesterolemia and in vitro method for altering nuclear receptor activity. The compounds are also used to reduce cholesterol levels, modulates cholesterol metabolism, increase cholesterol efflux from cells and increase the expression of ATP-Binding Cassette (ABC1) in the cells. The compounds are also useful for the treatment of macrovascular and macrovascular diseases e.g. nephropathy, neuropathy, retinopathy, hypertension, cerebrovascular disease, coronary heart disease, type 1 diabetes, type 2 diabetes, hyperlipidemia, Cushing syndrome, hypothyroidism, renal failure. The present sequence represents a human LXR beta polypeptide

Sequence 460 AA;

Alignment Scores:	2.39e-91	Length:	460
Pred. No.:	1215.50	Matches:	247
Score:	99.20%	Conservative:	0
Percent Similarity:	99.20%	Mismatches:	1
Best Local Similarity:	99.20%	Indels:	2
Query Match:	88.46%	Gaps:	1
DB:	7		

US-10-712-629B-2 (1-746) x ABR82544 (1-460)

1	Qy	AAGATTCCGGAAACAGCAGCAGAGAGTCA	60
169	Db	LysileArgLys---GlnGlnGlnSerGlnSerGlnSerProValGlyPro	187
61	Qy	CAGGGCAGCAGCAGCTCAGCCTCTGGGGCTTCCCTGTGTGGATCTCAGGCAGGC	120
188	Db	GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly	207
121	Qy	AGCCAGGGCTCCGGGGAAGCGAGGGTGTCCAGCTTAACAGCGGCTCAAGAACTAATGATC	180
208	Db	SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle	227
181	Qy	CAGCAGTGTGGGGCCCACTGCAGTGCACAAACGCTCCTTCTCCGACAGGCCCAA	240
228	Db	GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys	247
241	Qy	GTACGGCCTCGCCCTCGGGCGCCAGACCCAGTCCCGAGATGCCCGCCAGCAACGCTTT	300
248	Db	ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe	267
301	Qy	GCCCACTTCAGGAGCTGGCATCATCTCAGTCCAGGAGATCGTGGACTTCCTCAAGCAA	360
268	Db	AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln	287
361	Qy	GTGCCTGTGTTTCTGCACGTGGCGCGGAGGACACAGATCGCCCTCTCTGAAGCATCCACT	420
288	Db	ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr	307
421	Qy	ATCGAGATCATGCTGTAGACAGCCAGGCGCTACAACCGACGACAGAGTGTATCAC	480
308	Db	IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr	327

This Page Blank (Page 4)

|||||
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCGGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACAGCTCTCTTCCGACCAAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGGCCCTGGCGGAGAGCCCGAGTCCCGAGATCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 268
QY 301 GCCCACTTCACGAGGTGGCCATCATCTCAGTCCAGAGATCTGTGACTTCGTGACTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCTGGTTTCTCGAGTGGCGGAGAGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATATGCTGTAGAGACAGCGCGCTACACACAGACAGACAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCACCGTGCAGCGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCATCGCGCGCTGGCGCTTGGACGCGT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGGCCAGCGGCCCAAGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCGTGGAGGGTTCAGCAGCGCCCTAGCTGGAGCGCTGTCTTACACGCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCGCGCAGGACGAGTGGCG 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 2
159354
orphan nuclear receptor OR-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004
C;Accession: I59354; I59264
R;Teboul, M.; Enmark, E.; Li, Q.; Wikstrom, A.C.; Pello-Huikko, M.; Gustafsson, J.A.
Proc. Natl. Acad. Sci. U.S.A. 92, 2096-2100, 1995
A;Title: OR-1, a member of the nuclear receptor superfamily that interacts with the 9-ci
A;Reference number: I59354; MUID:95199298; PMID:7892230
A;Accession: I59354
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-446 <RES>
A;Cross-references: UNIPROT:Q62755; UNIPARC:UPT0000130508; EMBL:U20389; NID:9665941; PID
A;Experimental source: Sprague-Dawley, hepatic
R;Song, C.; Kokontis, J.M.; Hiipakka, R.A.; Liao, S.
Proc. Natl. Acad. Sci. U.S.A. 91, 10809-10813, 1994
A;Title: Ubiquitous receptor: a receptor that modulates gene activation by retinoic acid
A;Reference number: I59264; MUID:95062154; PMID:7971966
A;Accession: I59264
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-32, 'V', 34-51, 55-218, 'V', 220-446 <RE2>
A;Cross-references: UNIPARC:UPT00001458B3; EMBL:U14533; NID:9565661; PIDN:AAA52361.1; PI
A;Experimental source: vagina

C;Keywords: DNA binding; zinc finger
F;76-366/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 1-77e-77 Length: 446
Score: 1101.00 Matches: 225
Percent Similarity: 92.77% Conservatives: 6
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 80.13% Indels: 7
DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x I59354 (1-446)

QY 1 AAGATTCCGAACACAGCAGCAGAGTACAGTCACAGTCGAGTCACCTGCTGGGGCCG 60
Db 160 LysIleGlnLysGlnGlnGln-----GlnProProProThr 173
QY 61 CAGGCGCAGCAGCAGCTCAGCCCTCGGGCTCGGGCTTCCCTGGTGGATCTCAGGCGAGC 120
Db 174 GluProAlaSerGlySerSerAlaArgProAlaIleSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGGCTCCGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
Db 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACAGCTCTCTTCCGACCAAGCCCAAA 240
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCCCTGGGCCCTGGCGGAGAGCCCGAGTCCCGAGATCCCGCCAGCAACGCTTT 300
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
QY 301 GCCCACTTCACGAGGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTCCTGGTTTCTCGAGTGGCGGAGGAGACAGATCGCCCTCTCTTGAAGGCATCCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATCTGCTAGACAGCAGCGCGCTACACACAGCAGACAGATGTATCACC 480
Db 294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAGGACGAGTTCACCGCTGCAGCGCTGCAGGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCATCGGGCGCTGGCGCTGGAGCGCT 599
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 353
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGGCCAGACCGGCCCAAGTCGAGGAG 659
Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGGGTTCAGCAGCGCCCTAGCTGGAGCGCTGTCTTCTTACAGCGCGC 719
Db 374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY 720 ATCAAGAGCGCGCAGGACGAGTGGCG 746
Db 394 IleLysArgProGlnAspGlnLeuArg 402

RESULT 3

I49021

retinoid X receptor interacting protein No.15 - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004

C;Accession: I49021

R;Seol, W.; Choi, H.S.; Moore, D.D.

Mol. Endocrinol. 9, 72-85, 1995

A;Title: Isolation of proteins that interact specifically with the retinoid X receptor:
A;Reference number: A57664; MUID:95280959; PMID:7760852

A;Accession: I49021
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: mRNA
A;Residues: 1-446 <RES>
C;Genetics:
A;Gene: Rip15
C;Keywords: zinc finger
F;76-366/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 2,76e-77 Length: 446
Score: 1098.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.95% Indels: 2
DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x I49021 (1-446)

```
QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 155 GlnileArglyslsArglleGlnlys---GlnGlnGlnGlnProProProSer 173
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 CAGGGCAGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGCGATCTGAGGCAGGC 120
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 174 GluProAlaAlaSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetile 213
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 181 CAGCAGTTGGTGGCGGCCCACTGACAGTGCACAAACAGCTCTCTCCGACAGCCCAAA 240
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 241 GTCACGCCCTGGGCCCTGGGCGCAGACCCGAGTCCGAGATGCCGCCAGCAACGCTTT 300
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 254 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 273
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 361 GTGCCCTGGTTCTCAGCTGGCGCGGAGGACCATCGCCCTCTCTGAAGGCATCCACT 420
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 421 ATCGAGATCATGCTGCTAGAGCAGCGCGCTCAACACGAGACAGAGTGATCACC 480
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 294 IleGluIleWetLeuLeuGlnThrAlaArgArgTyraAsnHisGluThrGluCysIleThr 313
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 481 TTCTTGAG-GACTTACACGACGAGCGACTTCCACCGTCGAGCGCTGAGGTGGAG 539
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 314 PheLeuLysAspPheThrTyrrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 333
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 540 TTCATCAACCCCNATCTTCAGTCTCGCGGCGATCGCGGCGTGGCCCTGACGACGCT 599
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 334 PheileAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 353
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCCAACGTCAGGAG 659
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 354 GluTyraLeuLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 660 CCGGGCGCGGTGGAGCGTTGACGACGCCCTACGTGGAGCGCTGCTGTCTACAGCGCG 719
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 374 ProSerArgValGluAlaLeuGlnGlnProTyrrValGluAlaLeuLeuSertyrThrArg 393
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 720 ATCAGAGCGCGCAGGACGAGTGGCG 746
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 394 IleLysArgProGlnAspGlnLeuArg 402
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 4

A56043

steroid hormone receptor-like protein RLD-1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 31-Dec-2004

C;Accession: A56043

R;Apfel, R.; Benbrook, D.; Lernhardt, E.; Ortiz, M.A.; Salbert, G.; Pfehl, M.

Mol. Cell. Biol. 14, 7025-7035, 1994

A;Title: A novel orphan receptor specific for a subset of thyroid hormone-responsive el

A;Reference number: A56043; MUID:95021230; PMID:7935418

A;Accession: A56043

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-445 <APP>

A;Cross-references: UNIPROT:Q62685; UNIPARC:UPI000013050D; GB:U11685; NID:9555751; PIDN

A;Note: authors translated the codon GAG for residue 73 as Ser

C;Keywords: steroid hormone receptor; zinc finger

F;94-365/Domain: erba transforming protein homology <ERBA>

Alignment Scores:

Pred. No.: 7,46e-48 Length: 445
Score: 718.50 Matches: 151
Percent Similarity: 75.40% Conservative: 36
Best Local Similarity: 60.89% Mismatches: 38
Query Match: 52.29% Indels: 24
DB: 2 Gaps: 2

US-10-712-629B-2 (1-746) x A56043 (1-445)

```
QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 176 LysLeuLysArgGlnGlnGlnGlnAlaGlnAlaThrSerValSerPro----- 192
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTGGGGCTTCCCTGCTGGATCTGAGGCAGGC 120
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 193 -----ArgValSerSerProProGlnValLeuPro----- 202
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 203 -----GlnLeuSerProGlnLeuGlyMetile 212
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 181 CAGCAGTTGGTGGCGGCCCAACTGAGTGCACAAACGCTCTCTTCCGACACGCCCAAA 240
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 213 GluLysLeuValAlaAlaGlnGlnCysAsnArgArgSerPheSerAspArgLeuArg 232
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 241 GTCACGCCCTGGGCCCTGGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 233 ValThrProTrpProIleAlaProAspProGlnSerArgGluAlaArgGlnArgPhe 252
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGAGCTTCGCTAAGCAA 360
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 253 AlaHisPheThrGluLeuAlaIleValSerValGlnGluIleValAspPheAlaLysGln 272
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 361 GTGCCCTGGTTCTCCTCAGCTGGCGCGGAGGACCATCGCCCTCTCTGAAGGCATCCACT 420
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 273 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLeuLysThrSerAla 292
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 421 ATCGAGATCATGCTGCTAGAGACAGCCAGCGGCTCAACACGAGACAGAGTGATCACC 480
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 293 IleGluValMetLeuLeuGluThrSerArgArgTyraAsnProGlySerGluSerIleThr 312
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 481 TTCTTGAG-GACTTACCTACAGCAAGNACGACTTCCACCGTCGAGCGCTGAGGTGGAG 539
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 313 PheLeuLysAspPheSerTyrrAsnArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 332
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 540 TTCATCAACCCCNATCTTCTGAGTCTCGGGCGCATCGCGGCGTGGCCCTGGACGACGCT 599
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 333 PheileAsnProIlePheGluPheSerArgSerMetAsnGlnLeuGlnLeuAsnAspAla 352
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 600 GAGTAGCCCTGCTCATCGCCATCAACATCTTCTCGGCGCAGCCGCCCAACGTCAGGAG 659
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 353 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 372
   ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

```
QY 660 CCGGGCCGCTGGAGCGTTGCAGACGCGCTACAGCCCTACGTGGAGCGCTGCTGCTCTACAGCGC 719
Db ::::::::::::::::::::
373 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 392
QY 720 ATCAAGAGCGCGCAGGACGACGCTG 743
Db ::::::::::::::::::::
393 IleAsnHisProHisAspArgLeu 400

RESULT 5
I38975
nuclear orphan receptor LXR-alpha - human
C:Species: Homo sapiens (man)
C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 05-Oct-2004
C:Accession: I38975
R:Willy, P.J.; Umesono, K.; Ong, E.S.; Evans, R.M.; Heyman, R.A.; Mangelsdorf, D.J.
Genes Dev. 9, 1033-1045, 1995
A:Title: LXR, a nuclear receptor that defines a distinct retinoid response pathway.
A:Reference number: I38975; MUID:95262897; PMID:7744246
A:Accession: I38975
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-447 <ERBA>
A:Cross-references: UNIPROT:Q13133; UNIPARC:UPI00000529F2; EMBL:U22662; NID:9726512; PID
F:96-367/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 1,29e-46 Length: 447
Score: 702.50 Matches: 149
Percent Similarity: 74.19% Conservative: 35
Best Local Similarity: 60.08% Mismatches: 41
Query Match: 51.13% Indels: 24
DB: 2 Gaps: 2

US-10-712-629B-2 (1-746) x I38975 (1-447)

QY 1 AAGATTCCGAAACAGCAGCAGCAGGAGTCAAGTCAAGTTCGAGTCGAGTCACCTGTGGGCGG 60
Db ::::::::::::::::::::
178 LysLeuLysArgGlnGluGlnGluGlnAlaHisAlaThrSerLeuProPro----- 194
QY 61 CAGGGCAGCAGCTCAGCTCAGCTTGGGCTGGGGCTTCCCTCGTGGATCTGAGGCGGC 120
Db ::::::::::::::::::::
195 -----ArgArgSerSerProGlnIleLeuPro----- 204
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCAGTCAAGCGGCTCAAGAACTAATGATC 180
Db ::::::::::::::::::::
205 -----GlnLeuSerProGlnGlnLeuGlyMetIle 214
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTCAACAAACGCTCTTCTCCGACCCAGCCCAA 240
Db ::::::::::::::::::::
215 GluLysLeuValAlaAlaGlnGlnGlnCysAsnArgArgSerPheSerAspArgLeuArg 234
QY 241 GTCAGCGCTGGCCCTGGGCGCAGACCCCGAGTCCCGAGATCCCGCCGACGACGCTTT 300
Db ::::::::::::::::::::
235 ValThrProTrpProMetAlaProAspProHisSerArgGluAlaArgGlnGlnArgPhe 254
QY 301 GCCCATTACAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db ::::::::::::::::::::
255 AlaHisPheThrGluLeuAlaIleValSerValGlnGluLeuValAspPheAlaLysGln 274
QY 361 GTCGCTGGTTTCTGCACTGGGCGGAGACAGATCGCCCTCTGAAAGGCATCCACT 420
Db ::::::::::::::::::::
275 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLeuLysThrSerAla 294
QY 421 ATCGAGATCATGCTGTAGAGACAGCCGCGCTACACACGACAGACAGAGTGTATCAC 480
Db ::::::::::::::::::::
295 IleGluValMetLeuLeuGluThrSerArgArgTyrAsnProGlySerGluSerIleThr 314
QY 481 TTCTTGAG-GACTTCACAGCAAGCAGCTTCCACCGCTGAGCGCTGCGAGTGGAG 539
Db ::::::::::::::::::::
315 PheLeuLysAspPheSerTyrAsnArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 334
```

```
QY 540 TTCATCAACCCCATCTTCAGTTCTCGCGGGCCATCGCGGGCTGGGCTGGACGACGCT 599
Db ::::::::::::::::::::
335 PheIleAsnProIlePheGluPheSerArgAlaMetAsnGluLeuGluLeuAsnAspAla 354
QY 600 GAGTACGCGCTCTCATCGCCATCAACATCTTCTCGCGGCGAGCCGCCCAACCTGGCAGGAG 659
Db ::::::::::::::::::::
355 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 374
QY 660 CCGGGCCGCTGGAGCGTTGCAGACGCGCTACGTGGAGCGCTGCTGCTCTACAGCGC 719
Db ::::::::::::::::::::
375 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 394
QY 720 ATCAAGAGCGCGCAGGACGACGCTG 743
Db ::::::::::::::::::::
395 IleHisProHisAspArgLeu 402

RESULT 6
A56590
ecdysteroid receptor homolog cEcRH - midge (Chironomus tentans)
C:Species: Chironomus tentans
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 31-Dec-2004
C:Accession: A56590
R:Imhof, M.O.; Rusconi, S.; Lezzi, M.
Insect Biochem. Mol. Biol. 23, 115-124, 1993
A:Title: Cloning of a Chironomus tentans cDNA encoding a protein (cEcRH) homologous to
A:Reference number: A56590; MUID:93250857; PMID:8485513
A:Accession: A56590
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-536 <IMH>
A:Cross-references: UNIPROT:P49882; UNIPARC:UPI0000050DEA; GB:S60739; NID:9385893; PIDN
C:Note: sequence extracted from NCBI backbone (NCBIN:132124, NCBI:132127)
C:Keywords: DNA binding; steroid hormone receptor; zinc finger
F:113-432/Domain: erba transforming protein homology <ERBA>

Alignment Scores:
Pred. No.: 5,13e-15 Length: 536
Score: 294.50 Matches: 74
Percent Similarity: 60.11% Conservative: 39
Best Local Similarity: 39.36% Mismatches: 70
Query Match: 21.43% Indels: 6
DB: 2 Gaps: 3

US-10-712-629B-2 (1-746) x A56590 (1-536)

QY 151 CAGCTAACAGCGCTCAAGAACTAATGATCCAGAGTTCGTGGCGGCCCAACTGCAGTGC 210
Db ::::::::::::::::::::
274 GlnLeuThrAlaAsnGlnValAlaValIleTyrLysLeuIleTyrTyrGlnAspGlyTyr 293
QY 211 AACAAAGCTCTCTTCGACCGACCGCCAAAGTCACGCCCTGGCCCTGGCGGCGAGACCCC 270
Db ::::::::::::::::::::
294 GluGlnProSerGluGluAspLeuLysArgIleThrThrGluLeuGluGluGluAsp 313
QY 271 CAGTCCCGAGATGCCCGCCAGCAACGCTTTCGCCACTTCACGAGAGTGGCCATCATCTCA 330
Db ::::::::::::::::::::
314 GlnGluHisGluAla-----AsnPheArgTyrIleThrGluValThrIleLeuThr 330
QY 331 GTCACGAGATCGTGGACTTCGCTAAGCAAGTGCCTGTTCTCTGACGTGGCGCCGGAG 390
Db ::::::::::::::::::::
331 ValGlnLeuIleValGluPheAlaLysGlyLeuProAlaPheIleLysIleProGlnGlu 350
QY 391 GACGAGATGCCCTCTGAAGGCATCCATATCGAGATCATGCTGTAGACAGACGACG 450
Db ::::::::::::::::::::
351 AspGlnIleThrLeuLeuLysAlaCysSerSerGluValMetMetLeuArgMetAlaArg 370
QY 451 CGCTACAACACGACAGACAGATGATCACCTCTTCAGGAGCTTCACC-TACAGCAAGGAC 509
Db ::::::::::::::::::::
371 ArgTyrAspHisAspSerAspSerIleLeuPheAlaAsnAsnThrAlaTyrThrLysGln 390
QY 510 GACTTCCACGCTGAGCGCTTCAGGTGGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGG 569
Db ::::::::::::::::::::
391 ThrTyrGlnLeuAlaGlyMetGlu--GluThrIleAspAspLeuLeuHisPheCysArg 409
```


Db 397 IieProLeuLeuProAspGluIleLeuAla-----LysCysGlnAlaArgAsnIle 413
QY 226 TCCGACGACGCCAAAGTCACGCCCTGGCCCTGGCGCA----- 264
Db 414 -----ProSerLeuThrTyrAsnGlnLeuAlaValIleTyrLysLeuIleTrpTyr 430
QY 265 -----GACCCCCAGTCCCGAGATGCCGCCAG----- 291
Db 431 GlnAspGlyTyrGluGlnProSerGluGluAspLeuArgArgIleMetSerGlnProAsp 450
QY 292 -----CAACGCTTTGCCACTTCACGAGCTGGCCATCTCA 330
Db 451 GluAsnGluSerGlnThrAspValSerPheArgHisIleThrGluIleThrIleLeuThr 470
QY 331 GTCCAGGAGATCGTGACTTCGTAAAGCAAGTCCCTGGTTCTCGCAGCTGGCGCCGGAG 390
Db 471 ValGlnLeuIleValGluPheAlaLysGlyLeuProAlaPheThrLysIleProGlnGlu 490
QY 391 GACCAGATCGCCCTCTCGAAGGCATCCACTATCGAGATCATGCTGTAGAGACGACGAG 450
Db 491 AspGlnIleThrLeuLeuLysAlaCysSerSerGluValMetMetLeuArgMetAlaArg 510
QY 451 CGGTACAACACACAGACAGAGTGTATCACCTTCTTGAGGACTTC-ACCTACACACAGGAC 509
Db 511 ArgTyrAspHisSerSerAspSerIlePhePheAlaAsnAsnArgSerTyrThrArgAsp 530
QY 510 GACTTCCACCGTGCAGCGCTGCAGGTGGAGTTCATCAACCCCATCTTCGAGTTCTCGCGG 569
Db 531 SerTyrLysMetAlaGlyMet---AlaAspAsnIleGluAspLeuLeuHisPheCysArg 549
QY 570 GCCATCGCGCGGTGGCCCTGGACGACGCTGAGTACGCTGTCTCATCGCCATCAACATC 629
Db 550 GlnMetPheSerMetLysValAspAsnValGluTyrAlaLeuLeuThrAlaIleValIle 569
QY 630 TTCTCGGCCACCGGCCCAACGTCGACGACCGCGCGCGCGCTGAGGCGTTCGACGACCCC 689
Db 570 PheSer---AspArgProGlyLeuGluLysAlaGlnLeuValGluAlaIleGlnSerTyr 588
QY 690 TACGTGGAGCGCTGCTGCTCTAC 713
Db 589 TyrIleAspThrLeuArgIleTyr 596

RESULT 9

JC7229
vitamin D receptor subtype a - Paralichthys olivaceus
C;Species: Paralichthys olivaceus
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 05-Oct-2004
C;Accession: JC7229
R;Suzuki, T.; Suzuki, N.; Srivastava, A.S.; Kurokawa, T.
Biochem. Biophys. Res. Commun. 270, 40-45, 2000
A;Title: Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder
A;Reference number: JC7229

A;Accession: JC7229
A;Molecule type: mRNA
A;Residues: 1-420 <SUZ>
A;Cross-references: UNIPROT:Q9IB73; UNIPARC:UPI00000FC24F; DDBJ:AB037673
A;Experimental source: intestine
C;Comment: This receptor is an important factor in calcium homeostasis and bone formation
C;Genetics:
A;Gene: vdr-a
C;Superfamily: Vitamin D3 receptor (VDR)
C;Keywords: bone; calcium transport; DNA binding; hormone receptor; intestine; vitamin D

Alignment Scores:
Pred. No.: 3.12e-13 Length: 420
Score: 271.50 Matches: 76
Percent Similarity: 44.80% Conservative: 36
Best Local Similarity: 30.40% Mismatches: 85
Query Match: 19.76% Indels: 54
DB: 2 Gaps: 4

US-10-712-629B-2 (1-746) x JC7229 (1-420)

QY 151 CAGCTAACAGCGGCTCAAGAACTATGATCCAGCAGTGGTGGCGGCCCACTGCAGTGC 210
Db 123 ArgLeuAsnGluGlnAlaArgMetIleSerSerLeuValGlnAlaHisIleThr 142
QY 211 ACAACAACGCTCTTCTCCGACGACGCCAAAGTCACGCCCTGGCGCGCAGACCCC 270
Db 143 TyrAspAlaSerTyrSerAspPheSerArgPheArgPro---ProValargGluGlyPro 161
QY 271 CAGTCCCGCA-----GATCCCCCAGCAA 294
Db 162 ValThrArgSerAlaSerArgAlaAlaSerLeuHisSerLeuSerAspAlaSerSerAsp 181
QY 295 CCGCTTT----- 300
Db 182 SerPheAsnHisSerProGluSerValAspThrLysMetAsnPheSerAsnLeuLeuMet 201
QY 300 ----- 300
Db 202 MetTyrGlnAspGlyAlaSerSerProAspSerSerGluGluAsnThrLysLeuSerMet 221
QY 301 ---GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAG 357
Db 222 LeuProHisLeuAlaAspLeuValSerTyrSerIleGlnLysValIleGlyPheAlaLys 241
QY 358 CAAGTGCCTGTTCTCGCAGCTGGCGCGGAGGACCATCGCTCCCTCCTGAAGGCATCC 417
Db 242 MetIleProGlyPheArgAspLeuThrAlaGluAspGlnIleAlaLeuLeuLysSerSer 261
QY 418 ACTATCGAGATCATGCTGTAGACAGCGCGCTACACACACACAGACAGTGTATC 477
Db 262 AlaIleGluIleLeuMetLeuArgSerAsnGlnSerPheSerLeuGluAspMetSerTrp 281
QY 478 ACCTTCTT---GAGGACTTCACACAGACGACTTCACCGTGCAGCGCTGCAG 533
Db 282 SerCysGlyGlyProAspPheLysTyrCysIleAsnAspValThrLysAlaGlyHisThr 301
QY 534 GTGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATCGCGCGCTGGCCCTGGAC 593
Db 302 LeuGluLeuLeuGluProLeuValLysPheGlnValGlyLeuLysLysLeuAsnLeuHis 321
QY 594 GAGCGTGAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCCAGCGGCCCAACGTG 653
Db 322 GluGluGluHisValLeuLeuMetGlyIleCysLeuLeuSerProAspArgProGlyVal 341
QY 654 CAGGACCGCGCGCGCTGGAGCGCTTCAGACGACCCCTACGTGGAGCGCTGTCTCTCTAC 713
Db 342 GlnAspHisAlaArgValGluGlnLeuGlnAspArgLeuProGluAlaLeuGlnAlaTyr 361
QY 714 ACCGCGATCAAGAGCGCGGACGACACGCTG 743
Db 362 IleArgIleAsnHisProGlyGlyArgLeu 371

RESULT 10

A31761

1,25-dihydroxyvitamin D-3 receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Oct-2004
C;Accession: A31761; A31367
R;Burmeister, J.K.; Wiese, R.J.; Maeda, N.; Deluca, H.F.
Proc. Natl. Acad. Sci. U.S.A. 85, 9499-9502, 1988
A;Title: Structure and regulation of the rat 1,25-dihydroxyvitamin D-3 receptor.
A;Reference number: A31761; MUID:89071726; PMID:2849110
A;Accession: A31761
A;Molecule type: mRNA

A;Residues: 1-423 <BUR>

A;Cross-references: UNIPROT:P13053; UNIPARC:UPI00001382C3; GB:J04147; GB:J03630; NID:92

R;Burmeister, J.K.; Maeda, N.; Deluca, H.F.

Proc. Natl. Acad. Sci. U.S.A. 85, 1005-1009, 1988

A;Title: Isolation and expression of rat 1,25-dihydroxyvitamin D-3 receptor cDNA.

A;Reference number: A31367; MUID:88124963; PMID:2829212

A;Accession: A31367

A;Molecule type: mRNA

A;Residues: 'R', 58-423 <BU2>

the TCG Bank (uspio)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspro)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:13:37 ; Search time 220 Seconds
(without alignments)
4784.765 Million cell updates/sec

Title: US-10-712-629B-2
Perfect score: 1374
Sequence: 1 aagattcgaaacagcagca.....ggccgagcagcagctgcgc 746

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USFTO.spool_p/US10712629/runat_25112005_143036_10996/app_query.fasta_1.903
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZEX=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10712629 @CNC 1 1 466 @runat_25112005_143036_10996 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -JONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	461	1 NRIH2_HUMAN	P55055 homo sapien
2	1231	89.6	461	2 Q510Y1_HUMAN	Q510Y1 homo sapien
3	1231	89.6	461	2 Q61BU6_HUMAN	Q61BU6 homo sapien
4	1189	86.5	460	2 Q68CY8_HUMAN	Q68CY8 homo sapien
5	1147.5	83.5	452	2 Q58D83_BOVIN	Q58D83 bos taurus
6	1144.5	83.3	455	1 NRIH2_BOVIN	Q58D86 bos taurus
7	1144.5	83.3	455	2 Q58BIN1_BOVIN	Q5BIN1 bos taurus
8	1144.5	83.3	498	2 Q58CP4_BOVIN	Q58CP4 bos taurus
9	1101	80.1	446	1 NRIH2_RAT	Q62755 mus musculus
10	1098.5	79.9	446	1 NRIH2_MOUSE	Q62755 mus musculus
11	1095.5	79.7	443	2 Q8BP65_MOUSE	Q8BP65 mus musculus
12	1028	74.8	351	2 Q5REL4_PONPY	Q5REL4 pongo pygma
13	885	63.0	441	2 Q6GNA2_XENLA	Q6GMA2 xenopus lae
14	731	53.2	412	2 Q5GA56_BRARE	Q5GA56 brachydanio
15	721.5	52.5	445	1 NRIH3_MOUSE	Q920Y9 mus musculus
16	719.5	52.4	445	2 Q91X41_MOUSE	Q91X41 mus musculus

17	719.5	52.4	447	2	Q4TU03_PIG	Q4TU03 sus scrofa
18	718.5	52.3	445	1	NRIH3_RAT	Q62685 rattus norv
19	718.5	52.3	445	2	Q51035_RAT	Q51035 rattus norv
20	712.5	51.9	447	2	Q5E9B6_BOVIN	Q5E9B6 bos taurus
21	707.5	51.5	402	2	Q81W13_HUMAN	Q81W13 homo sapien
22	704.5	51.3	447	1	NRIH3_HUMAN	Q13133 homo sapien
23	704.5	51.3	447	2	Q5U0N9_HUMAN	Q5U0N9 homo sapien
24	702.5	51.1	409	2	Q8JHU1_CHICK	Q8JHU1 gallus gall
25	690.5	50.3	409	2	Q8AXU8_CHICK	Q8AXU8 gallus gall
26	499	36.3	509	2	Q4SHF7_TETNG	Q4SHF7 tetradodon n
27	472.5	34.4	557	2	Q4H386_CIOIN	Q4H386 ciona intes
28	400.5	29.1	444	2	Q44336_9ACAR	Q44336 amblyomma a
29	400.5	29.1	560	2	Q44337_9ACAR	Q44337 amblyomma a
30	400.5	29.1	570	2	Q44338_9ACAR	Q44338 amblyomma a
31	331.5	24.1	328	2	Q5GCP7_GECIA	Q5GCP7 gecarcinus
32	331.5	24.1	518	2	Q76246_UCAPU	Q76246 uca pugilac
33	325.5	23.7	488	2	Q4W6C9_LEPDE	Q4W6C9 leptinotars
34	325.5	23.7	565	2	Q4W6D0_LEPDE	Q4W6D0 leptinotars
35	321	23.4	541	2	Q97095_LOCMI	Q97095 locusta mig
36	312.5	22.7	336	2	Q6R1B4_CARMA	Q6R1B4 carcinus ma
37	307	22.3	491	2	Q02035_TENMO	Q02035 tenebrio mo
38	299	21.8	503	2	Q6P7H5_XENLA	Q6P7H5 xenopus lae
39	299	21.8	513	2	Q8JU28_XENLA	Q8JU28 xenopus lae
40	297.5	21.7	530	2	Q8JU27_XENLA	Q8JU27 xenopus lae
41	294.5	21.4	536	1	ECR_CHITE	P49882 chironomus
42	291.5	21.2	556	1	ECR_MANSE	P49883 manduca sex
43	289.5	21.1	412	2	Q7PVB2_ANOGA	Q7PVB2 anopheles g
44	289.5	21.1	513	2	Q77240_CHOFU	Q77240 choristoneu
45	289.5	21.1	541	2	Q77255_CHOFU	Q77255 choristoneu

ALIGNMENTS

RESULT 1

RESULI I					
AD	NRIH2 HUMAN	STANDARD;	PRT;	461	AA.
AC	P55055; Q12970;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	13-SEP-2005 (Rel. 48, Last annotation update)				
DE	Oxytetracycline receptor LXR-beta (Liver X receptor beta) (Nuclear orphan				
DE	receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (Nuclear				
DE	receptor NER)				
GN	Name=NRIH2; Synonyms=LXRB, NER, UNR;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	TISSUE=Osteosarcoma;				
RX	MEDLINE=95011628; PubMed=7926814; DOI=10.1016/0378-1119(94)90080-9;				
RA	Shinar D.M., Endo N., Rutledge S.J., Vogel R., Rodan G.A., Schmidt A.;				
RT	NER, a new member of the gene family encoding the human steroid				
RT	hormone nuclear receptor.";				
RL	Gene 147:273-276(1994).				
RP	[2]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].				
RC	TISSUE=Blood, Placenta, and Uterus;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Shenmen C.M., Schuler G.D.,				
RA	Klausner R.D., Collins F.S., Wagner L., Schaefer C.F., Bhat N.K.,				
RA	Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fabey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				


```
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_NBD_1; 1.
DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 461 AA; 51087 MW; 697E3D9F9BC5C0BE CRC64;

Alignment Scores:
Pred. No.: 4,24e-80 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservativeness: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x Q61BU6_HUMAN (1-461)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysglnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCCCTGGGGCTTCCCTGTGTGATCTGAGGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaGlnGlnLeuMetIle 228
QY 181 CAGCAGTTGTGGGGCCCACTGCAGTGCAACAAAGCTCTCTCCGACCCAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCAGCGCTCCGCTGGCGGAGACCCCGAGTCCGAGATGCCGCCAGCAAGCGCTT 300
Db 249 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnAlaArgPhe 268
QY 301 GCCACTTCAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTGGTTTCTGAGCTGGCGGAGGAGCAGATCGCCCTCTCGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGlyAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGAGACGACGAGCGCTACACACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCGAGCGCTCGAGGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTATCAACCCCATCTTCGAGTTCTCGGGGCCATGCGGGCGCTGGGGCTGGAGCAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGCTCATCGGCATCAACATCTTCTCGGCCGACCGCCCAACAGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGCTGGAGCGCTTGCACAGCCCTTACGTGGAGCGCTGTCTCTACACCGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAGAGGGCGCAGGACCGCTGGC 746
```

```
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 4
Q68CY8_HUMAN
ID Q68CY8_HUMAN PRELIMINARY; PRT; 460 AA.
AC Q68CY8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686D1580.
GN Name=DKFZp686D1580;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_
R1 NUCLEOTIDE SEQUENCE.
R2 TISSUE=Endometrium carcinoma cell line;
RG The German cDNA Consortium;
RA Ansoer W., Krieger S., Regiert T., Rittmuller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity). Receptor family.
CC -1- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; CR749648; CAH18442.1; -; mRNA.
DR SMR; Q68CY8; 220-459.
DR Ensembl; ENSG00000131408; Homo sapiens.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR001723; Strhmn_receptor.
DR InterPro; IPR000324; Vitd_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_NBD_1; 1.
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
KW Receptor; Transcription; Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 460 AA; 50401 MW; FD0063D952B23D94 CRC64;

Alignment Scores:
Pred. No.: 4,45e-77 Length: 460
Score: 1189.00 Matches: 247
Percent Similarity: 98.41% Conservativeness: 0
Best Local Similarity: 98.41% Mismatches: 1
Query Match: 86.54% Indels: 4
DB: 2 Gaps: 0

US-10-712-629B-2 (1-746) x Q68CY8_HUMAN (1-460)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LysileArglysglnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCCCTGGGGCTTCCCTGTGTGATCTGAGGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGGAAGCGAGGTGTCCAGCTAACAGCGGTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaGlnGlnLeuMetIle 228
QY 181 CAGCAGTTGTGGGGCCCACTGCAGTGCAACAAAGCTCTCTCCGACCCAGCCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCAGCGCTCCGCTGGCGGAGACCCCGAGTCCGAGATGCCGCCAGCAAGCGCTT 300
Db 249 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnAlaArgPhe 268
QY 301 GCCACTTCAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTGGTTTCTGAGCTGGCGGAGGAGCAGATCGCCCTCTCGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGlyAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGAGACGACGAGCGCTACACACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCGAGCGCTCGAGGTGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTATCAACCCCATCTTCGAGTTCTCGGGGCCATGCGGGCGCTGGGGCTGGAGCAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGCTCATCGGCATCAACATCTTCTCGGCCGACCGCCCAACAGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGCTGGAGCGCTTGCACAGCCCTTACGTGGAGCGCTGTCTCTACACCGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAGAGGGCGCAGGACCGCTGGC 746
```

```
QY 181 CAGCAGTTGGTGGCGGCCCACTGCAAGTCAACAAACGCTCTCTCCGACGACCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCCCTGGCGCCCTGGCGGAGACCCCACTCCGAGATCCCGCCAGCAACGCTTT 300
Db 249 ValThrProThrProLysGlyAlaAspProGlnSerArgAspAlaArgGlnAlaArgPhe 268
QY 301 GCCCACTTACAGGAGTGGCCATCATCTCAGTCCAGGAGATCTGCACTTCCGTAGCCAA 360
Db 269 AlaHisPheThrGlnLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCTCTGTTTCTTCAGCTGGCGGAGGACCATCGCTCCCTGAGGCGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGAGACAGCGCGCTCAACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-GACTTACCTACAGAGGAGCTTCCACCGTCAGCGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrThrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCGCATCGCGGCTGGCGCTGGACGAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 368
QY 600 GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGGCGGACCGGCCCAACGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArg-ProThrCysArg-A 388
QY 660 CCGGGCGCGGTGAGCGGTTCAGACGACCCCTAGCTGAGCGCGCTGTCTCTACAGCGCG 719
Db 388 IaglyProArgGlyGlyValAlaAlaLeuArgGlyGlyAlaAlaValLeuHisAlaH 408
QY 720 ATCAAGAGCGCGGACGACGAGTGGCG 746
Db 408 IeGlnGluAlaAla-AspGlnLeuArg 416
```

RESULT 5

```
Q58D83_BOVIN
ID Q58D83_BOVIN PRELIMINARY; PRT; 452 AA.
AC Q58D83;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Nuclear receptor subfamily 1, group H, member 2.
GN Names=NR1H2;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Poolled;
RX MEDLINE=21180013; PubMed=11282978; DOI=10.1101/gr.170101;
RA Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
RA Casas E., Wray J.E., White J., Cho J., Fahrenkrug S.C., Bennett G.L.,
RA Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
RA Pertea G., Holt I., Karamycheva S., Liang F., Quackenbush J.,
RA Keele J.W.;
RT "Sequence evaluation of four pooled-tissue normalized bovine cdna
libraries and construction of a gene index for cattle."
RL Genome Res. 11:626-630(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Poolled;
RA Harhay G.P., Sonstegard T.S., Van Tassel C.P., Clawson M.L.,
RA Heaton M.P., Keele J.W., Snelling W.M., Weidmann R.T., Smith T.P.L.;
RA "Sequencing and analysis of Bos taurus full-length insert cdna
```

```
RT clones.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; BT021714; AAX4561.1; -; mRNA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003355; P:regulation of transcription factor activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0006350; P:transcription; IEA.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 452 AA; 49662 MW; 8E9B32E29C57825A CRC64;
```

Alignment Scores:

```
Alignment Scores: 4.3e-74 Length: 452
Pred. No.: 1147.50 Matches: 235
Score: 1147.50 Conservative: 2
Percent Similarity: 94.80% Mismatches: 10
Best Local Similarity: 94.00% Indels: 4
Query Match: 83.52% Gaps: 2
DB: 2
```

US-10-712-629b-2 (1-746) x Q58D83_BOVIN (1-452)

```
QY 1 AAGATTTCGGAACAGCAGCAGGAGTCAAGTCAAGTCCAGTCCAGTCCCTGTGGGCGG 60
Db 161 LysIleArgLysGlnGlnGlnGln-----GlnGlnGlnSerSerProThrGlyPro 178
QY 61 CAGGCGCAGCAGCAGCTCA---GCCTCTGGGCGCTGGGGCTTCCCTGGTGTATCTGAGCA 117
Db 179 GlyValSerSerSerSerProAlaSerGlyProGlyAlaSerProGlyGlySerAspGly 198
QY 118 GCGAGCCAGCGCTCCGGGGAAGCGAGGCTGTCCAGCTAACAGCGGCTCAAGAACTAATG 177
Db 199 GlyGlyGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMet 218
QY 178 ATCCAGCAGCTTGTGGCGGCCCAACTGCAAGTGCACAAACGCTCTCTCCGACGAGCCC 237
Db 219 IleGlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnPro 238
QY 238 AAAGTCACGCCCTTGGCGCCCTGGCGGAGGACCCCAAGTCCCGAGATCCCGCCAGCAACGC 297
Db 239 LysValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArg 258
QY 298 TTTGCCCACTTACAGGAGTGGCCATCATCTCAGTCCAGGAGATCTGTGACTTTCGCTAAG 357
Db 259 PheAlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLys 278
QY 358 CAAGTGCCTGTTTCTCTGAGTGGCGGCGGAGGACCATCGCTCCCTCTGAAGGCGATCC 417
Db 279 GlnValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSer 298
QY 418 ACTATCGAGATCATGTCTAGACAGCAGCGCGCTACAACCCAGACAGAGTGTATC 477
Db 299 ThrIleGluIleMetLeuLeuGluThrAlaArgTyrAsnHisGluThrGluCysIle 318
QY 478 ACTTCTTTCAG-GACTTACCTACAGCAAGGACGACTTCCAGCTGCAGCGCTGCAGGTG 536
Db 319 ThrPheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnVal 338
QY 537 GAGTTTCATCAACCCCATCTTCAGTTCCTCGCGGCGCATCGCGGCTGGGCGCTGGAGCAC 596
Db 339 GluPheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAsp 358
QY 597 GCTGAGTACGCCCTGTCTCATGCCATCAACATCTTCTCGGCGGACCGGCCCAACGTCGAG 656
Db 359 AlaGluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGln 378
QY 657 GAGCGCGCGCGGTGAGCGGTTTGAGCGAGCGCTAGTGGAGCGCTGTCTGTCTCTACAGG 716
Db 379 GluProSerArgValGluAlaLeuGlnGlnProTyrValAspAlaLeuLeuSerTyrThr 398
```


CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.

DR EMBL; BT021903; AAX46750.1; -; mRNA.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 498 AA; 54139 MW; 29CED3E4E6F4156 CRC64;

Alignment Scores:
 Pred. No.: 7,09e-74 Length: 498
 Score: 1144.50 Matches: 235
 Percent Similarity: 94.80% Conservative: 2
 Best Local Similarity: 94.00% Mismatches: 10
 Query Match: 83.30% Indels: 4
 DB: Gaps: 2

US-10-712-629B-2 (1-746) x Q58CP4_BOVIN (1-498)

QY 1 AGATTTCGAAACAGCAGCAGGAGTCACTACAGTCAGTCGAGTCACCTGTGGGGCGG 60
 DB 164 LysileArglyseGlnGlnGln-----GlnGlnGlnSerSerProThrGlyPro 181
 QY 61 CAGGGCAGCAGCAGCTCA---GCCTCTGGGCGCTTCCCTCGTGATCTGAGGCA 117
 DB 182 GlyValSerSerSerProAlaSerGlyProGlyAlaSerProGlyGlySerAspGly 201
 QY 118 GGCAGCAGGGCTCCGGGAGCGAGGGTGTCCAGCTAACACGGCTCAAGAACTAATG 177
 DB 202 GlyGlyGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaGlnGlnLeuMet 221
 QY 178 ATCAGCAGTGTGTGGGGCCCACTGCAGTGAACAAAGCTCTCTCCGACCGCC 237
 DB 222 IleGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnPro 241
 QY 238 AAGTTCAGCGCTGGCGCTGGCGGAGACCCCGAGTCCCGAGATCCCGCCAGCAACGC 297
 DB 242 LysValThrProThrProLeuGlyAlaAspProGlnSerArgAlaArgGlnArg 261
 QY 298 TTGCCCATTACGAGCTGGGCATCATCTAGTCCAGAGATCGTGGACTTCGCTAAG 357
 DB 262 PheAlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLys 281
 QY 358 CAAGTGCCTGTTCTCTGAGCTGGCGGAGGACAGATCCGCTCTGAGGATCC 417
 DB 282 GlnValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSer 301
 QY 418 ACTATCGAGATCATGCTGTAGACAGCCAGCGCTTACACCAACGAGACAGAGTGTATC 477
 DB 302 ThrIleGluLeuMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIle 321
 QY 478 ACCTTCTTGAAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCGAGCGCTGAGGTG 536
 DB 322 ThrPheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnVal 341
 QY 537 GAGTTCATCAACCCCATCTTCGAGTCTCCGGGGCCATCGGGCTGGCGCTGGAGCAG 596
 DB 342 GluPheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAsp 361
 QY 597 GCTGAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCGCAGCCGCCCAAGTGCAG 656
 DB 362 AlaGluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGln 381
 QY 657 GAGCCGGCGCGTGGAGGGCTTCCAGCAGCCCTACCTGAGGGCGCTGCTGCTCTACAG 716
 DB 382 GluProSerArgValGluAlaLeuGlnGlnProTyrValAspAlaLeuLeuSerTyrThr 401
 QY 717 CGCATCAAGAGCGCGAGGACCACTCGCG 746
 DB 402 ArgIleLysArgProGlnAspGlnLeuArg 411

RESULT 9
 NR1H2 RAT STANDARD; . PRT; 446 AA.
 AC Q62755; Q62694;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Oysterols receptor LXR-beta (liver X receptor beta) (Nuclear orphan
 DE receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (UR)
 DE (Orphan nuclear receptor OR-1).
 GN Name=Nr1h2; Synonyms=Lxrb;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]_TaxID=10116;
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=95199298; PubMed=7892230;
 RA Teboul M., Enmark E., Li Q., Wikstrom A.C., Pelto-Huikko M.,
 Gustafsson J.-A.;
 RT "OR-1, a member of the nuclear receptor superfamily that interacts
 RT with the 9-cis-retinoic acid receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:2096-2100(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sprague-Dawley; TISSUE=Vagina;
 RX MEDLINE=95062154; PubMed=7971966;
 RA Song C., Kokontis J.M., Hiipakka R.A., Liao S.;
 RT "Ubiquitous receptor: a receptor that modulates gene activation by
 RT retinoic acid and thyroid hormone receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10809-10813(1994).
 CC -!- FUNCTION: Orphan receptor. Binds preferentially to double-stranded
 CC oligonucleotide direct repeats having the consensus half-site
 CC sequence 5'-AGTCA-3' and 4-nt spacing (DR-4).
 CC -!- SUBUNIT: Forms a heterodimer with RXR.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
 CC subfamily.
 CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; U20389; AAA69522.1; -; mRNA.
 CC EMBL; U14533; AAA52361.1; -; mRNA.
 DR PIR; I59354; I59354.
 DR HSP; P55055; I59354.
 DR SNR; O62755; 205-445.
 DR TRANSFAC; T04451; -;
 DR Ensembl; ENSRNOG0000019812; Rattus norvegicus.
 DR RGD; 61906; Nr1h2.
 DR InterPro; IPR001628; Hrmn rcpt DNA bd.
 DR InterPro; IPR000536; Hrmn rcpt lig bd.
 DR InterPro; IPR001723; Stdhrmn receptor.
 DR InterPro; IPR000324; Vitd_receptor.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMNER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00350; VITAMINDR.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 DR PROSITE; PS51030; NUCLEAR_REC_DBD_2; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.

```
FT  DNA_BIND  75  152  Nuclear receptor.
FT  ZN_FING   78  98   NR C4-type.
FT  ZN_FING   116 140  NR C4-type.
FT  REGION    216 446  Ligand-binding (Potential).
FT  COMPTIAS  162 168  Poly-Gln.
FT  COMPTIAS  169 172  Poly-Pro.
FT  CONFLICT  33  33  G -> V (in Ref. 2).
FT  CONFLICT  52  54  Missing (in Ref. 2).
FT  CONFLICT  219 219  A -> V (in Ref. 2).
SQ  SEQUENCE  446 AA; 49736 MW; 13DF6DC2F0F5FA4D CRC64;

Alignment Scores:
Pred. No.: 9.5e-71
Score: 1101.00
Percent Similarity: 92.77%
Best Local Similarity: 90.36%
Query Match: 80.13%
DB: 1
Gaps: 1

US-10-712-629B-2 (1-746) x NR1H2_RAT (1-446)
QY 1 AAGATTCCGAAACAGCAGCAGCAGGAGTACAGTCACAGTCGCGAGTCACCTGTGGGGCGG 60
Db 160 LysileinylsGlnGlnGln-----GlnProProProProThr 173
QY 61 CAGGCGCAGCAGCAGCAGCTCTGGGCTTGGGGCTTCCCTGGTGATCTGAGGAGCG 120
Db 174 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCAGGCTCCGGGAGCGAGGGTGTCACAGTAAACAGCGCTCAAGAACCTAATGATC 180
Db 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetile 213
QY 181 CAGCAGTTGGTGGCGCCCACTGCAGTCGACCAACAGCGCTCTTCCGACGAGCCCAA 240
Db 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCTTGGCCCTTGGCGCAGACCCCGAGTCCCGAGATCCCGCCAGCAGCAACGCTT 300
Db 234 ValThrProThrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCATTCAGCAGTGGCATCATCTCAGTCCAGAGATCTGTGACTCGCTAAGCAAA 360
Db 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTGCCCTGGTTCTCGCAGCTGGCGGAGGACAGATCGCCCTCTGAGGAGCATCACT 420
Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTGATAGACAGCAGCGCGCTACAACACAGACAGAGTGTATCACC 480
Db 294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACTACAGCAGGAGCAGCTTCCACCGTGCAGCGCTGCGAGTGGAG 539
Db 314 PheLeuLysAspPheThrTyrrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCATCGCGGCTGGGCTTGGACGAGCT 599
Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTAGCGCTGCTCATCGCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
Db 354 GluTyrrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGCGGTTCAGAGCCCTACGTGAGGCGCTGTCTCTACACGCCG 719
Db 374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrrThrArg 393
QY 720 ATCAGAGCGCGCAGCAGCAGCTGCCG 746
Db 394 IleLysArgProGlnAspGlnLeuArg 402
```

```
RESULT 10
NR1H2_MOUSE
ID NR1H2_MOUSE STANDARD; PRT; 446 AA.
AC Q60644
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Oxyterols receptor LXR-beta (Liver X receptor beta) (Nuclear orphan
DE receptor LXR-beta) (Ubiquitously-expressed nuclear receptor) (Retinoid
DE X receptor interacting protein No.15).
GN Names=Nr1h2; Synonyms=Lxrb, Rlp15, Unr, Unr2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=95280959; PubMed=7760852; DOI=10.1210/me.9.1.72;
RA Seol W., Choi H.S., Moore D.D.;
RT "Isolation of proteins that interact specifically with the retinoid X
RT receptor: two novel orphan receptors.";
RL Mol. Endocrinol. 9:72-85(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129/SvJ; TISSUE=Liver;
RX MEDLINE=20156373; PubMed=10675617; DOI=10.1016/S0378-1119(99)00555-7;
RA Alberti S., Steffensen K.R., Gustafsson J.-A.;
RT "Structural characterisation of the mouse nuclear oxysterol receptor
RT genes LXRalpha and LXRbeta.";
RL Gene 243:93-103(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Orphan receptor. Binds preferentially to double-stranded
CC oligonucleotide direct repeats having the consensus half-site
CC sequence 5'-AGGTCA-3' and 4-nt spacing (DR-4).
CC -!- SUBUNIT: Forms a heterodimer with RXR.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR1
CC subfamily.
CC -!- SIMILARITY: Contains 1 nuclear receptor DNA-binding domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC EMBL; U09419; AAC52164.1; -; mRNA.
DR
```

DR EMBL; AJ132602; CAB51924.1; -; Genomic_DNA.
 DR BMBL; BC066025; AAH66025.1; -; mRNA.
 DR PIR; I49021; I49021.
 DR HSSP; P55055; 1P8D.
 DR SMR; Q60644; 205-445.
 DR TRANSFAC; T04467; -.
 DR Ensembl; ENSMUSG0000060601; Mus musculus.
 DR MGI; MGI:1352463; Nr1h2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IDA.
 DR GO; GO:0003700; F:transcription factor activity; IDA.
 DR GO; GO:004255; P:cellular lipid metabolism; IDA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
 DR InterPro; IPR001628; Hrmn_rcpt_DNA_bd.
 DR InterPro; IPR000536; Hrmn_rcpt_lig_bd.
 DR InterPro; IPR001723; Strhmn_receptor.
 DR InterPro; IPR000324; VitD_receptor.
 DR Pfam; PF00104; Hormone_recep; 1.
 DR Pfam; PF00105; zf-C4; 1.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR00047; STROIDFINGER.
 DR PRINTS; PR00350; VITAMINDR.
 DR ProDom; PD000035; Znf_C4steroid; 1.
 DR SMART; SM00430; HOL1; 1.
 DR SMART; SM00399; Znf_C4; 1.
 DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
 DR PROSITE; PS1030; NUCLEAR_REC_DBD_2; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger;
 FT DNA_BIND 75 152 Nuclear receptor.
 FT ZN_FING 78 98 NR C4-type.
 FT ZN_FING 116 140 NR C4-type.
 FT REGION 216 446 Ligand-binding (Potential).
 FT COMPTAS 162 168 Poly-Gln.
 FT COMPTAS 169 172 Poly-Pro.
 SQ SEQUENCE 446 AA; 49720 MW; 73153B635302C9DF CRC64;

 Alignment Scores:
 Pred. No.: 1.44e-70 Length: 446
 Score: 1098.50 Matches: 224
 Percent Similarity: 93.57% Conservative: 9
 Best Local Similarity: 89.96% Mismatches: 15
 Query Match: 79.95% Indels: 2
 DB: 1 Gaps: 1

 US-10-712-629b-2 (1-746) x NR1H2_MOUSE (1-446)
 QY 1 AAGATTGGAAACAGCAGCAGCAGGAGTACAGTCACAGTCGGCAGTCACCTGTGGGCGG 60
 DB 155 GlnIleArgLysLysArgIleGlnLys--GlnGlnGlnGlnGlnProProProSer 173
 QY 61 CAGGGCAGCAGCAGCAGCAGCTCTGGGCGCTTCCCTGGTGATCTGAGGCAGGC 120
 DB 174 GluProAlaAlaSerSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
 QY 121 AGCCAGGGCTCCGGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
 DB 194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
 QY 181 CAGCAGTGTGGGGCGGCCCACTGCAGTCGACAAAGCTCTCTCCGACCCAGCCAAA 240
 DB 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
 QY 241 GTACGCGCTGGCGCCCTGGGCGCAGACCCCGAGATGCCCGCCAGCAACGCTTT 300
 DB 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
 QY 301 GCCCACTTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTTAACAA 360
 DB 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
 QY 361 GTGCTGTGTTCTCGAGCTGGGCGGAGCAGATCGCCCTCCTGAGGCACTCCACT 420

Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
 QY 421 ATCGAGATCATGCTCTAGACACGCCAGGCGCTTCAACCCAGACAGAGTGTATCACC 480
 Db 294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
 QY 481 TTCCTTGAG-GACTTCACCTACAGCAGCAGCTTCCACCGTCAGGCGCTGCAGGTGGAG 539
 Db 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
 QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGCGCTGGACGCGT 599
 Db 334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 353
 QY 600 GAGTAGCGCTCTCATCGCCATCAACATCTTCTCGCGCGCAGCGGCCCACTGCAGGAG 659
 Db 354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
 QY 660 CCGGGCGGCTGGAGCGTTCGAGCAGCGCTTACGTGAGGCGCTGCTGTCTACACGCGC 719
 Db 374 ProSerArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
 QY 720 ATCAAGAGCGCGCAGCAGCAGCTGCGC 746
 Db 394 IleLysArgProGlnAspGlnLeuArg 402

 RESULT 11
 Q8BP65_MOUSE PRELIMINARY; PRT; 443 AA.
 AC Q8BP65_MOUSE PRELIMINARY; PRT; 443 AA.
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched
 DE library, clone:5730478C17 product:nuclear receptor subfamily 1, group
 DE H, member 2, full insert sequence.
 DE Name=Nr1h2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,
 RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bareh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).

[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs";
RN Nature 420:563-573 (2002).
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RL prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630 (2000).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771 (2000).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saichon H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Taya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AK077620; BAC36906.1; -; mRNA.
DR HSSP; P55055; 1PQ6.
DR SMR; Q89P65; 202-442.
DR MGI; MGI:1352463; Nr1h2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0004879; F:ligand-dependent nuclear receptor activity; IDA.
DR GO; GO:0003700; F:transcription factor activity; IDA.
DR GO; GO:0044255; P:cellular lipid metabolism; IDA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR001723; Stdhrm_receptor.
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRODHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 443 AA; 49394 MW; 362DCB2D21034F29 CRC64;

Alignment Scores:
Pred. No.: 2,36e-70 Length: 443
Score: 1095.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.73% Indels: 2
DB: 2 Gaps: 1
US-10-712-629B-2 (1-746) x Q8BP65_MOUSE (1-443)
QY 1 AAGATTCCGAAACGAGCAGCAGGAGTCCACATCAGTCGAGTCGACCTGTGGGCGC 60
Db 152 GlnileArglysyysArgilleGlnLys---GlnGlnGlnGlnProProProSer 170
QY 61 CAGGCGCAGCAGCAGCTCAGCCTCTGGGCGCTTCCCTGGTGGATCTGAGGCGAGC 120
Db 171 GluProAlaAlaSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 190
QY 121 AGCAGGCGCTCCGGGGAAGCGAGGCTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 191 SerGlnGlySerGlyGluGlyGluGlyGluGlyGluGlyGluGluGluGluGluGlu 210
QY 181 CAGCAGTTGGTGGCGCCCAACTGCAAGTGCACAAACAGCTCTCTCCGACCCAGCCAAA 240
Db 211 GlnGlnLeuValAlaAlaGlnLeuGlnCysAenLysArgSerPheSerAepGlnProLys 230
QY 241 GTCACCCCTGGCGCCCTGGGCGGAGAGCCCGCAGTCCGAGATCCCGCCAGCAACGCTTT 300
Db 231 ValThrProTrpProLeuGlyAlaAepProGlnSerArgAepAlaArgGlnGlnArgPhe 250
QY 301 GCCCACTTCACGAGCTGGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAGCA 360
Db 251 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 270
QY 361 GTGCTCTGTTTCTCAGCTGGCGGAGGAGCAGATCGCCCTCTCTGAGGATCCACT 420
Db 271 ValProGlyPheLeuGlnLeuGlyArgGluAepGlnGlnAlaLeuLeuLysAlaSerThr 290
QY 421 ATCGAGATCATGCTGTAGACAGCAGCCGCTACACCCAGCAGAGATGATATCACC 480
Db 291 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 310
QY 481 TTCTTGAG-GACTTTCACCTACAGCAGGAGCAGCTCCACCGTCAGCGCTGAGGTCGAG 539
Db 311 PheLeuLysAepPheThrTyrSerLysAepAepPheHisArgAlaGlyLeuGlnValGlu 330
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCGCATCGCGGCTGGGCGCTGGACGCGCT 599
Db 331 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAepAla 350
QY 600 GAGTACGCGCTCTCATCCCATCAACATCTTCTCGCGCGAGCCGCGCCAACTGCAGGAG 659
Db 351 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAepArgProAsnValGlnGlu 370
QY 660 CCGGCGCGGTGGAGCGTTCGAGAGCCCTACGTCAGCGAGCGCTGAGGCGCGCTGCTACACGCGC 719
Db 371 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 390
QY 720 ATCAAGAGCGCCGAGCAGCAGCTGCGC 746
Db 391 IleLysArgProGlnAepGlnLeuArg 399
RESULT 12
QSREL4_PONPY PRELIMINARY; PRT; 351 AA.
ID QSREL4_PONPY
AC QSREL4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp468A0622.
GN Name=DKFZp468A0622;
OS Pongo pygmaeus (Orangutan).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OX Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Heart;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; CR857510; CAH89793.1; -; mRNA.
DR SMR; Q5REL4; 140-350.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0003708; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000536; Hrmn_recept_lig.
DR InterPro; IPR003078; Rtnoid_receptor.
DR InterPro; IPR001723; Stdhmn_receptor.
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR01292; RETNOICACIDR.
DR PRINTS; PR00398; STRDHORMONR.
DR PRINTS; PR00047; STDRIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_REC_DBD_1; 1.
KW DNA-binding; Hypothetical protein; Nuclear protein;
KW Receptor; Transcription; Transcription regulation; Zinc; zinc-finger.
SQ SEQUENCE 351 AA; 39475 MW; 175C907978A0247C CRC64;

Alignment Scores:
Pred. No.: 1.68e-65 Length: 351
Score: 1028.00 Matches: 214
Percent Similarity: 86.35% Conservative: 1
Best Local Similarity: 85.94% Mismatches: 4
Query Match: 74.82% Indels: 31
DB: 2 Gaps: 1

US-10-712-629b-2 (1-746) x Q5REL4_PONPY (1-351)

QY 1 AAGATTGGAAACAGACGACGAGGAGTTCACAGTCGAGTCACCTGTGGGCGG 60
DB 89 LysileargylsGlnGlnGlnGlnSerGlnSerGlnSerGlnSerProAlaGlyPro 108
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCGCTTGGGCGCTTCCCTGGTGGATCTGAGCGAGC 120
DB 109 GlnGlySerCysSerThrSerGlyProGlyAlaSerProGlyGlySerGlyAlaGly 128
QY 121 AGCAGGGCTCCGGGAGGCGAGGCTGTCAGCTAACAGGGCTCAAGAACTAATGATC 180
DB 129 SerGlnGlySerGlyGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetile 148
QY 181 CAGCAGTTGGTGGGCGCCCAACTGACGTCAACAAAGCTCTTCTCCGACCGAGCCCAAA 240
DB 149 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnlysaargserPheSerAspGlnProlys 168
QY 241 GTACGCGCTTGGCGCTTGGGCGGACAGCCCGAGTCCCGAGATGCCCGCAGCAACGCTTT 300
DB 169 ValThr----- 170
QY 301 GCCCACTTCAGGAGCTGGCCATCATCTCAGTCAGGAGATCGTGACTTCGCTTAAGCAA 360
|||||

Db 171 -----GluIleValaspPheAlaIysGln 178
QY 361 GTGCTGTGGTTCTCAGCTGGCGGAGGAGCAGATCGCCTCTCTGAAGCATCCACT 420
|||||
Db 179 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 198
QY 421 ATCGAGATCATCTCTAGACACACCCAGCGCTCAACACCAGACAGACAGTGTATCACC 480
|||||
Db 199 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 218
QY 481 TTCTTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCGAGGCTCGCAGGTGGAG 539
|||||
Db 219 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 238
QY 540 TTTCATCAACCCCATCTCTCAGTTCTCGCGGCGCATCGCGCGCTGGCGTCGACGACGCT 599
|||||
Db 239 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 258
QY 600 GAGTACGGCTCTCTCATGCCATCAACATCTTCTCGGCGGACCGCCCAACGTGCAGGAG 659
|||||
Db 259 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 278
QY 660 CCGGCGCGCTGCGAGCGTTGCGAGCGCTACGTGGAGCGCTGTCTCTACACGCGC 719
|||||
Db 279 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 298
QY 720 ATCAAGAGCGCGCAGGACGAGCTGCGC 746
|||||
Db 299 IleLysArgProGlnAspGlnLeuArg 307

RESULT 13
Q6GMA2 XENLA
ID Q6GMA2 XENLA PRELIMINARY; PRT; 441 AA.
AC Q6GMA2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC81974 protein.
GN Names=MGC81974;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC Tissue=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,

Richardson P.;
 "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative";
 Dev. Dyn. 225:384-391 (2002).
 [3].
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Gerhard D.S.;
 Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
 DR EMBL: BC074169; AAH74169.1; -, mRNA.
 DR SMR: Q6GMA2; 200-440.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0004884; F:ecdysteroid hormone receptor activity; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0005496; F:steroid binding; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR InterPro: IPR003069; Ecdystd_receptor.
 DR InterPro: IPR000536; Hrmn_recept_lig.
 DR InterPro: IPR001723; Strhrm_receptor.
 DR InterPro: IPR003324; Vitd_receptor.
 DR InterPro: IPR001628; Znf_C4steroid.
 DR Pfam: PF00104; Hormone_recep; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PRO1283; ECDYSTEROIDE.
 DR PRINTS: PRO0398; STRDHORMONER.
 DR PRINTS: PRO0047; STROIDFINGER.
 DR ProDom: PD000035; Znf_C4steroid.
 DR ProDom: PD000035; Znf_C4steroid; 1.
 DR SMART: SM00430; HOLI; 1.
 DR SMART: SM00399; ZNF_C4; 1.
 DR PROSITE: PS00031; NUCLEAR_REC_DBD_1; 1.
 DR DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 SQ SEQUENCE 441 AA; 50229 MW; AF7FD8B8F11741CB CRC64;

Alignment Scores:
 Pred. No.: 9e-54 Length: 441
 Score: 865.00 Matches: 184
 Percent Similarity: 77.82% Conservative: 9
 Best Local Similarity: 74.19% Mismatches: 35
 Query Match: 62.95% Indels: 21
 DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x Q6GMA2_XENLA (1-441)

QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTTCACAGTCACAGTCGACGTCACCTGTGGGCGC 60
 Db 169 LyslleArlgLyGlnGlnAspValThrArgSerSerAlaLeuValProPro 188
 QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGGCGCTGGGGCTTCCCTGGTGATCTGAGGCAGGC 120
 Db 189 SerProCysMetLeuSerGln----- 195
 QY 121 ACCCAGGGCTCCGGGAAGCGCGGTGTCAGCTAACAGCGCTCAAGAACTAATGATC 180
 Db 196 -----GluValValGlnLeuThrProGlnGlnGlnLysMetlle 208
 QY 181 CAGCAGTTCGGTGGCGCCCAACTGCGAGTGCACAAACGCTCTCTCCGACAGCCCAAA 240
 Db 209 GluGlnLeuValSerAlaGlnGlnGlnCysAsnLysArgSerPheSerAspGlnProLys 228
 QY 241 GTCAGCCCTGCGCCCTGGGCGCAGACCCCAAGTCCCGAGATGCGCCGACGAACTTT 300
 Db 229 ValThrProTrpProGlyThrAspProAsnSerArgGluAlaArgGlnGlnArgPhe 248
 QY 301 GCCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCCTTAAGCAA 360
 Db 249 AlaHisPheThrGluLeuAlaIlelleSerValGlnGlnValGluValAspPheAlaLysGln 268

QY 361 GTGCCTGGTTCTTCAGCAGTCGGCGGAGGAGCAGATCGCCCTCTCTGAGGATCCACT 420
 Db 269 ValProGlyPheLeuGluLeuSerArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 288
 QY 421 ATCGAGATCATGCTGTAGACAGCCAGCGCTACAAACACGAGACAGAGTGTATCACC 480
 Db 289 IleGluIleMetLeuLeuGluThrAlaArgArgTrpAsnHisGluThrGluCysIleThr 308
 QY 481 TTCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCGAGCGCTGCGAGTGGAG 539
 Db 309 PheLeuLysAspPheThrTrpSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 328
 QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGGGCATCGCGCGCTCGGACGACGCT 599
 Db 329 PheIleAsnProIlePheGluPheSerArgGlyMetArgGlnMetGlnLeuAspAspAla 348
 QY 600 GAGTACGCGCTGCTCATGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
 Db 349 GluTrpAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValLeuAsn 368
 QY 660 CCGGGCGCGTGGAGCGTTGCGAGCGCTCAGTGGAGCGCTGCTGCTCTACACGCGC 719
 Db 369 HisGlnHisValGluAsnLeuGlnLeuProTrpValGluAlaLeuHisSerTrpThrArg 388
 QY 720 ATCAAGAGCGCGCAGGACCGAGCTG 743
 Db 389 IleLysArgProGlnAspHisLeu 396

RESULT 14
 Q56A56 BRARE PRELIMINARY; PRT; 412 AA.
 AC Q56A56;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein zgc:113339.
 GN Name=zgc:113339;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Singapore local strain; TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smaluk D.E.,
 RA Scherch A., Schain J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Singapore local strain; TISSUE=Embryo;
 RL NIH MGC Project;
 RG Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.

```
DR ENBL; BC092160; AAH92160.1; -, mRNA.
```

```
DR GO: GO:0005634; C:nucleus; IEA.
```

```
DR GO: GO:0003707; F:steroid hormone receptor activity; IEA.
```

```
DR GO: GO:0004887; F:thyroid hormone receptor activity; IEA.
```

```
DR GO: GO:0003700; F:transcription factor activity; IEA.
```

```
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

```
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
```

```
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;
```

```
KW Receptor; Transcription; transcription regulation; Zinc; Zinc-finger.
```

```
SQ SEQUENCE      412 AA;  47679 MW;  30B5C4B8CD1AC86D CRC64;
```

```
Alignment Scores:
```

Pred. No.:	3,91e-44	Length:	412
Score:	731.00	Matches:	154
Percent Similarity:	74.19%	Conservative:	30
Best Local Similarity:	62.10%	Mismatches:	42
Query Match:	53.20%	Indels:	23
DB:	2	Gaps:	2

US-10-712-629B-2 (1-746) x Q56A56 BRARE (1-412)

Qy	1	AAGATTGGAAACAGCAGCAGCAGGAGTCA	CAGTCCAGTCCGAGTCACTCTGGGGCGG	60
Db	142	ArgLeuLybLysMetLysGlnGluGluThr	-----	153
Qy	61	CAGGGCAGCAGCAGCTCAGCCTCTGGGCT	CTCCCTGGTGGATCTCAGGCAGCG	120
Db	154	---AlaArgThrSerThrValAlaThrProSer	ProAlaPro-----	166
Qy	121	AGCCAGGGCTCCGGGGAAGCGAGGTCTCC	AGCTTAACAGCGGCTCAAGAACTAATGTC	180
Db	167	-----GluMetProProLeuAlaProGluGln	GluMetIle	179
Qy	181	CAGCAGTTGGTGGCGGCCAACTGCGAGTG	CAACAAACGCTCTCTTCGCACAGCGCCAAA	240
Db	180	GluIysLeuValAlaMetGlnIysGlnCysAsn	LysArgSerPheIleAspArgProLys	199
Qy	241	GTCAGCCCTCGCCCTCGGCGCAGACCCCA	CCAGTCCCGAGATCCGCCAGCAACGCTTT	300
Db	200	ValThrProIrrProGlnSerGlnAspProGln	AsnArgGluValArgGlnGlnArgPhe	219
Qy	301	GCCCACTTCAGCGAGCTGGCCATCATCTC	AGTCCAGGAGATCGTGGACTTCCTTAAGCAA	360
Db	220	AlaHisPheThrGluLeuAlaIleMetSerVal	GlnGluIleValAspPheAlaLysGln	239
Qy	361	GTGCTGTGTTCTCGACGCTGGCGCGGAGG	ACACAGATCGCCTCTGAAGCATCCACT	420
Db	240	LeuProGlyPheLeuGluLeuThrArgGluAsp	GlnIleAlaLeuLeuLysThrSerThr	259
Qy	421	ATCAGATCATGCTCTAGACACGCCAGGCGC	TACAACACAGACAGATGATCACCC	480
Db	260	IleGluIleMetLeuLeuGluThrSerArgTyr	AsnProAlaIleAspSerIleThr	279
Qy	481	TTCTTGAG-GACTTCACCTACAGCAGGACG	ACTTCCACCGTCAGGCCTCGCAGGTGGAG	539
Db	280	PheLeuLysAspPheThrTyrAsnLysGluAsp	PheAlaLysAlaGlyLeuGlnLeuGlu	299
Qy	540	TTCATCAACCCCATCTTCGAGTTCTCGGGG	CCATGCGCGCGCTGGCCTCGACACAGCT	599
Db	300	PheIleAsnProIlePheGluPheSerLysGlyMet	AsnAspLeuHisLeuAspGluAla	319
Qy	600	GAGTACGCCCTGCTCATCGCCATCAACATCT	TCTCGGCGCAGCGCCCAACGTGACGAG	659
Db	320	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSer	AlaAspArgProAsnValGlnAsp	339
Qy	660	CCGGCCCGCTGGGCGGCTTGACAGACGCCCT	TACGTGGAGCGCTGCTGCTTACACGCGC	719
Db	340	HisGluLeuValGluArgLeuGlnProTyrValAsp	AlaLeuHisSerTyrIleArg	359
Qy	720	ATCAAGAGCGCCGACGACCATCGTCTG	743	
Db	360	IleLysArgProAsnAspHisLeu	367	

RESULT 15

KW Transcription regulation; Zinc; Zinc-finger.
FT DNA_BIND 93 168 Nuclear receptor.
FT ZN_FING 96 116 NR C4-type.
FT ZN_FING 132 156 NR C4-type.
FT REGION 213 432 Ligand-binding (Potential).
FT CONFLICT 399 R -> P (in Ref. 2).
SQ SEQUENCE 445 AA; 50477 MW; C9A8DF38D935593 CRC64;

Alignment Scores:

Pred. No.:	1-89e-43	Length:	445
Score:	721.50	Matches:	152
Percent Similarity:	75.40%	Conservative:	35
Best Local Similarity:	61.29%	Mismatches:	38
Query Match:	52.51%	Indels:	24
DB:	1	Gaps:	2

US-10-712-629B-2 (1-746) x NR1H3_MOUSE (1-445)

```
QY 1 AGATTCCGAAACAGCAGCAGCAGGAGTCACATCAGTCGCACTGCGAGTCACCTGTGGGCGCG 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 176 LysLeuLysArgGlnGluGlnGlnAlaGlnAlaThrSerValSerPro----- 192

QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCGCTGGGCGCTTCCCTCGTGATCTGAGGCAGGC 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 193 -----ArgValSerSerProGlnValLeuPro----- 202

QY 121 AGCCAGGGCTCCGGGAAGCGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATC 180
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 -----GlnLeuSerProGlnLeuGlyMetIle 212

QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACGGCTCTCTCCGACCACCCAAA 240
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 213 GluLysLeuValAlaAlaGlnGlnCysAsnArgSerPheSerAspArgLeuArg 232

QY 241 GTCACGCCCTGGCGCCCTGGGCGGAGACCCCGAGTCCCGAGATGCCCGCAGCAACCGCTT 300
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 233 ValThrProTrpProIleAlaProAspProGlnSerArgGluAlaArgGlnArgPhe 252

QY 301 GCCCACTTCACGAGGTGCCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 253 AlaHisPheThrGluLeuAlaIleValSerValGlnGluIleValAspPheAlaLysGln 272

QY 361 GTGCTGTGTTCTCTGAGCTGGCGGAGGACAGATCGCCCTCTGAAGGCATCCACT 420
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 273 LeuProGlyPheLeuGlnLeuSerArgGluAspGlnIleAlaLeuLysThrSerAla 292

QY 421 ATCGAGATCATGCTGCTAGACAGCAGCGCGCTACAACACGACAGACAGTGTATCACC 480
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 293 IleGluValMetLeuLeuGluThrSerArgTyrAsnProGlySerGluSerIleThr 312

QY 481 TTCCTTGAG-GACTTACCTACAGCAAGGACGACTCCACCGTCGAGCGCTGCGAGTGGAG 539
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 313 PheLeuLysAspPheSerTyrAsnArgGluAspPheAlaLysAlaGlyLeuGlnValGlu 332

QY 540 TTCATCAACCCCATCTTTCAGATTCTCGCGGCGCATCGCGCGCTGGACACGCT 599
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 333 PheIleAsnProIlePheGluPheSerArgAlaMetAsnGluLeuGlnLeuAsnAspAla 352

QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCGGACCGGCGCCAAACGTGACGAG 659
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 353 GluPheAlaLeuLeuIleAlaIleSerIlePheSerAlaAspArgProAsnValGlnAsp 372

QY 660 CCGGGCGCGGTGGAGCGTTGCAGCAGCCCTAGCTGGAGGCGCTGCTCTACACGCGC 719
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 373 GlnLeuGlnValGluArgLeuGlnHisThrTyrValGluAlaLeuHisAlaTyrValSer 392

QY 720 ATCAGAGGCGCGCAGGACCGAGCTG 743
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 393 IleAsnHisProHisAspArgLeu 400
```

Search completed: November 25, 2005, 22:30:18
Job time : 231 secs

This Page Blank (Page 2)


```
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 288
QY 301 GCCACTTACAGAGCTGGCCATCATCTAGTCCAGGAGATCGTGACTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTCTGTTCTCTGAGCTGGCGGAGGAGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGAGACAGCCAGCGCTACAACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTCAG-GACTTACCTACAGCAAGGAGCACTTCCACGCTGAGCGCTCGAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGGGCCGACATCGCGCGCTCGGCGCTGAGGAG 600
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCGGACCGCGCCCAACGTGCGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGGTGGAGCGTTGACAGCAGCCCTAGTGGAGCGCTGCTGTCTACACGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCCCGCAGGACGAGTGGCG 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
```

RESULT 3

```
US-08-646-248-2
; Sequence 2, Application US/08646248
; Patent No. 5939322
; GENERAL INFORMATION:
; APPLICANT: Friedman, Eitan
; APPLICANT: Holloway, M. Katharine
; APPLICANT: Rodan, Gideon
; APPLICANT: Rutledge, Su Jane
; APPLICANT: Schmidt, Azriel
; APPLICANT: Vogel, Robert
; TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,248
; FILING DATE: 14-MAY-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330,283
; FILING DATE: 27-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Dolan, Catherine A.
; REGISTRATION NUMBER: 36,502
; REFERENCE/DOCKET NUMBER: 19327
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4283
```

```
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-646-248-2
Alignment Scores:
Pred. No.: 1,1e-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 1 Gaps: 0
```

US-10-712-629B-2 (1-746) x US-08-646-248-2 (1-461)

```
QY 1 AAGATTCGAAACAGCAGCAGGAGTCAAGTCAAGTCAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysIleArgLysGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTGGGGCTTCCCTGGTGGATCTGAGGCGAG 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAGGCGAGGTGTCCAGCTAACACAGCGGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGACGTGCAACAAACGCTCTCTTCCGACACGCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCAGCGCTGGCGCCCTGGGCGCAGACCCCGAGATGCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCACTTACAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTCTGTTCTCTGAGTGGCGGAGGAGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATGCTGTAGAGACAGCCAGCGCTACAACACAGACAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTCAG-GACTTACCTACAGCAAGGAGCACTTCCACGCTGAGCGCTCGAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGGGCCGACATCGCGCGCTCGGCGCTGAGGAG 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTGGCGGAGCGCGCCCAACGTGCGAGG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCCGGTGGAGCGTTGACAGCAGCCCTAGTGGAGCGCTGCTGTCTACACGCC 719
Db 389 ProGlyArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408
QY 720 ATCAAGAGCCCGCAGGACGAGTGGCG 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
```



```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-13931-2

Alignment Scores:
Pred. No.: 1.1e-104 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservatve: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 99.59% Indels: 1
DB: 8 Gaps: 0

US-10-712-629B-2 (1-746) x PCT-US95-13931-2 (1-461)

QY 1 AAGATTCGGAACAGCAGCAGGAGTCAAGTCAAGTCAGTCGAGTCACCTGTGGGGCGG 60
Db 169 LysileArglyGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188

QY 61 CAGGCGAGCAGCAGCTCAGCTCTGGCGCTGGGGCTTCCCTGTGTGATCTGAGGCGG 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208

QY 121 AGCCAGGGCTCCGGGGAAGCGAGGGGTCCAGCTAACAGCGGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228

QY 181 CAGCAGTTGTGGCGGCCCACTGCAGTGCACAAACAGCTCTTCTCCGACCCAGCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248

QY 241 GTACGCCCTTCCGGCGGAGAGCCAGCCAGTCCGAGATCCCGAGATCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268

QY 301 GCCCATTACCGAGCTGGGCGCATCATCTCAGTCCAGGAGATCTGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288

QY 361 GTCCCTGGTTTCTGAGCTGGCGGAGGACAGATCCCTCTCTGAAAGGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLysAlaSerThr 308

QY 421 ATCAGATCATGTCTAGACAGCAGCGCTACACCCAGTCCAGAGTGTATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGlnThrGluCysIleThr 328

QY 481 TTCTTGAG-CACCTTACCTACAGCAGGAGCAGCTTCCACCGTGCAGGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 348

QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCATCGCGCGCTGGCGCTGGACGAGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368

QY 600 GAGTAGCCCTGTCTATCGCCATCAACATCTTCTCGGCCAGCCGCCAACGTCGAGGAG 659
Db 369 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388

QY 660 CCGGGCGCGTGGAGGGTTGACAGAGCCCTACGTGGAGCGCTGTCTGTCTACAGCCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 408

QY 720 ATCAAGAGCGCGCAGGACCAAGCTGCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 6
US-09-949-016-7717
; Sequence 7717, Application US/09949016
```


Db 386 GluTyAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 405
Qy 660 CCGGGCGGCTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGCTGTCTACACGCGC 719
Db 406 ProGlyArgValGlnAlaLeuGlnProTyValGlnAlaLeuLeuSerTyThrArg 425
Qy 720 ATCAAGAGCGCGCAGGACCACTGCGC 746
Db 426 IleLysArgProGlnAspGlnLeuArg 434
RESULT 7
US-09-976-594-655
; Sequence 655, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; FILE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 655
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3721108CD1
US-09-976-594-655
Alignment Scores:
Pred. No.: 2,55e-104 Length: 461
Score: 1227.00 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 2
Query Match: 89.30% Indels: 1
DB: 2 Gaps: 0
US-10-712-629b-2 (1-746) x US-09-976-594-655 (1-461)
Qy 1 AAGATTCGGAACAGCAGCAGCAGGTTCACAGTCACAGTCGAGTCACCTGTGGGCGC 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
Qy 61 CAGGGCAGCAGCAGCTCAGCTCTTGGGCTTGGGCTTCCCTGTGGATCTGAGCGAGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
Qy 121 AGCAGGCTCCGGGGAAGCGAGGCTGTCAGCTACAGCGGTCAAGACTAATCATC 180
Db 209 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
Qy 181 CAGCAGTTGGTGGGCGGCCCACTCAGTGCAACAAACGCTCTCTCCGACACGCCAAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
Qy 241 GTACGCGCTGGCCCTTGGGCGCAGACCCCGAGATCCCGCCAGCAACAGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 268
Qy 301 GCCCACTTCAGGAGCTGGCCATCATCTCAGTCAGGAGATCGTGGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleSerValGlnGluIleValAspPheAlaLysGln 288
Qy 361 GTGCTGCTGTTCTTCAGCTGGGCGGAGGAGACCATCGCCCTCCTGAAGGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308

Qy 421 ATCGAGATCATCTCTAGACAGCAGCGGCTACACACGAGACAGAGTGATATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyAsnHisGluThrGluCysIleThr 328
Qy 481 TTCTTGAG-GACTTCACCTACAGCAGGACGACTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db 329 PheLeuLysAspPheThrTySerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 348
Qy 540 TTTCATCAACCCCATCTTCGAGTTCTCGGGCCATCGCGCGCTGGGCTTGACAGCGCT 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 368
Qy 600 GAGTACGCGCTCTCATCGCCATCAACATCTTCTCGGCGCAGCGGCCCACTGACGAGG 659
Db 369 GluTyAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
Qy 660 CCGGGCGGCTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGCTGTCTACACGCGC 719
Db 389 ProGlyArgValGlnAlaLeuGlnProTyValGluThrLeuLeuSerTyThrArg 408
Qy 720 ATCAAGAGCGCGCAGGACCACTGCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417
RESULT 8
US-08-342-411A-2
; Sequence 2, Application US/08342411A
; Patent No. 5639616
; GENERAL INFORMATION:
; APPLICANT: LIAO, Shuteung
; APPLICANT: SONG, Ching
; TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,411A
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KITCHELL, BARBARA S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: ARCD154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 460 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-342-411A-2
Alignment Scores:
Pred. No.: 5,94e-104 Length: 460
Score: 1223.00 Matches: 246
Percent Similarity: 99.20% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 2
Query Match: 89.01% Indels: 1
DB: 1 Gaps: 0


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-325-2

Alignment Scores:
Pred. No.: 9,39e-93 Length: 446
Score: 1101.00 Matches: 225
Percent Similarity: 92.77% Conservative: 6
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 80.13% Indels: 7
DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x US-09-909-325-2 (1-446)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
DB 160 LysileGlnLysGlnGlnGln-----GlnProProProThr 173
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
DB 174 GluProAlaSerGlySerSerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACACAGCGCTCAAGAACTAATGATC 180
DB 194 SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCACTGAGTCAGTCAACAAACGGCTCTTCCAGACGCCAAA 240
DB 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCCCTGGCGCTGGCGCAGACCCCGAGTCCGAGATGCCCGCCAGCAACGCTTT 300
DB 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
QY 301 GCCACTTCACGAGCTGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA 360
DB 254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAlaAspPheAlaLysGln 273
QY 361 GTGCTGTGTTCTCAGCTGGCGCGGAGGAGCAGATGCCCTCTCTGAAGGATCCACT 420
DB 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCGAGATCATGCTGTAGAGCAGCAGCGCGCTACACACAGCAGACAGTGTATCACC 480
DB 294 IleGluLeuLeuLeuGluThrAlaArgArgTyraAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACCGTCAGGCTCGCAGGTGGAG 539
DB 314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCAGATTCTCGGGGCCATCGGGCGCTGGGCTGGACGAGCT 599
DB 334 PheileAsnProIlePheGluPheSerArgAlaMetArgGluGlyLeuAspAspAla 353
QY 600 GAGTAGCCCTGCTCATCGCATCAACATCTTCTCGCGCGCAGCGCGCCCAACGTCAGGAG 659
DB 354 GluTyraLeuLeuLeuLeuAlaIleAenIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCCGGTGGAGCGGTTGACAGCAGCCCTAGCTGGAGGCGGTGTGTCTACACGCGC 719
DB 374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY 720 ATCAGAGCGCGCAGACACAGCTGGCGC 746
DB 394 IleLysArgProGlnAspGlnLeuArg 402

RESULT 12
US-09-909-326-2
; Sequence 2, Application US/09909326
; Patent No. 6617120
; GENERAL INFORMATION:
```

```

; APPLICANT: ENMARK, EVA
; GUSTAFSSON, JAN
; TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING
; TO THE NUCLEAR RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/909,326
; FILING DATE: 19-Jul-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/776,844
; FILING DATE: 1997-06-24
; APPLICATION NUMBER: UK 9413536.2
; FILING DATE: 16-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kegan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 00487.04029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-326-2

Alignment Scores:
Pred. No.: 9,39e-93 Length: 446
Score: 1101.00 Matches: 225
Percent Similarity: 92.77% Conservative: 6
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 80.13% Indels: 7
DB: 2 Gaps: 1

US-10-712-629B-2 (1-746) x US-09-909-326-2 (1-446)
QY 1 AAGATTTCGGAACAGCAGCAGCAGGAGTACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
DB 160 LysileGlnLysGlnGlnGln-----GlnProProProThr 173
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
DB 174 GluProAlaSerGlySerSerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACACAGCGCTCAAGAACTAATGATC 180
DB 194 SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCACTGAGTCAGTCAACAAACGGCTCTTCCAGACGCCAAA 240
DB 214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCCCTGGCGCTGGCGCAGACCCCGAGTCCGAGATGCCCGCCAGCAACGCTTT 300
DB 234 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
```

QY 301 GCCACTTCACGAGCTGGCCATCATCTCAGTCCAGAGATCGTGACTTCGTGAAGCAA 360
Db |||||PheThrGluLeuAlaIleIleSerValGlnGlnGluValAspPheAlaLysGln 273
QY 361 GTCCCTGGTTCTCGAGCTGGCCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db |||||PheLeuGlnLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCCAGATCATGCTGTAGACAGCCAGCGCTACAAACACAGACAGAGTGTATCACC 480
Db |||||IleGluLeuLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACGTCAGCCCTGCAGGTGGAG 539
Db |||||PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTTCATCAACCCCATCTCGAGTTCTCGCGGCCCATCGCGGGCTGGCCCTGGACGCGCT 599
Db |||||PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCCGACCGCCCAACGTCGACGAG 659
Db |||||GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGCGCTTGACAGCAGCCCTACGTGGAGCGCTGTGCTCTACAGCGC 719
Db |||||ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 393
QY 720 ATCAGAGCGCGGACGAGCTGGCC 746
Db |||||IleLysArgProGlnAspGlnLeuArg 402

RESULT 13

US-08-372-652-3
; Sequence 3, Application US/08372652
; Patent No. 5932699
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Seol, Wongi
; APPLICANT: Choi, Hweng-Sik
; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
; TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street, Suite 3100
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 13-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/246001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-372-652-3

Alignment Scores:

Pred. No.: 3e-92 Length: 446
Score: 1095.50 Matches: 223
Percent Similarity: 93.57% Conservative: 10
Best Local Similarity: 89.56% Mismatches: 15
Query Match: 79.73% Indels: 2
DB: 1 Gaps: 1

US-10-712-629B-2 (1-746) x US-08-372-652-3 (1-446)

QY 1 AAGATTCGGAACACAGCAGCAGAGTCAAGTCAAGTCAAGTCCAGTCACTGTGGGCGCG 60
Db |||||AArgLysArgIleGlnLys---GlnGlnGlnGlnGlnProProProSer 173
QY 61 CAGGGCAGCAGCAGCTCAGCCTCTGGGCTTCCCTGGTGGATCTGAGGCGAGGC 120
Db |||||GluProAlaIleSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGCTCCGGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db |||||SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAAGTCAACAAACGCTCTTCTCCGACCAAGCCCAA 240
Db |||||GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCACGCCCTGGGCCCTGGCGCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db |||||ValThrProIleProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCACTTCACGAGCTGGCCATCATCTCAGTCCAGAGATCGTGACTTCGTGAAGCAA 360
Db |||||AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTCCCTGGTTCTCGAGCTGGCCGGGAGGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db |||||ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCCAGATCATGCTGTAGACAGCAGCGCTACAAACACAGACAGAGTGTATCACC 480
Db |||||IleGluLeuLeuLeuGlnThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGACGACTTCCACGCTGAGCGCTTCGAGGTGGAG 539
Db |||||PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTTCATCAACCCCATCTCGAGTTCTCGCGGCCCATCGCGGGCTGGCCCTGGACGCGCT 599
Db |||||PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla 353
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGGCCGACCGCCCAACGTCGAGG 659
Db |||||GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGTGGAGCGCTTGACAGCAGCCCTACGTGGAGCGCTGTGCTCTACAGCGC 719
Db |||||ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuSerTyrThrArg 393
QY 720 ATCAGAGCGCGGACGAGCTGGCC 746
Db |||||IleLysArgProGlnAspGlnLeuArg 402

RESULT 14

PCT-US95-16311-3
; Sequence 3, Application PC/TUS9516311
; GENERAL INFORMATION:
; APPLICANT: Moore, David
; APPLICANT: Seol, Wongi
; APPLICANT: Choi, Hweng-Sik
; TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING

TITLE OF INVENTION: POLYPEPTIDES AND RELATED MOLECULES AND METHODS

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street, Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16311
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/372,652
FILING DATE: 13-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/246001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 446 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16311-3

Alignment Scores:
Pred. No.: 3e-92 Length: 446
Score: 1095.50 Matches: 223
Percent Similarity: 93.57% Conservative: 10
Best Local Similarity: 89.56% Mismatches: 15
Query Match: 79.73% Indels: 2
DB: 4 Gaps: 1
US-10-712-629B-2 (1-746) x PCT-US95-16311-3 (1-446)

QY 1 AGATTGCGAAACAGCAGCAGCAGGAGTCACTACAGTCGAGTCACCTGTGGGGCCG 60
Db :::::::::::::::::::: |||||
155 GlnIleArgIysIysArgIleGlnLys--GlnGlnGlnGlnGlnProProProSer 173
QY 61 CAGGGCAGCAGCAGCTCAGCTCTGGCCCTGGGGCTTCCCTGGTGATCTGAGGCAGGC 120
Db :::::::::::::::::::: |||||
174 GluProAlaIleSerSerSerGlyArgProAlaIleSerProGlyThrSerGluAlaSer 193
QY 121 AGCCAGGCTCCGGGGAAGCAGGAGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db :::::::::::::::::::: |||||
194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaIleGlnGluLeuMetIle 213
QY 181 CAGCAGTGTGTGGCGGCCCACTGCAGTGCACAAACAGCTCTCTCCGACACGCCCAA 240
Db :::::::::::::::::::: |||||
214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY 241 GTCAGCCCTGGCCCTGGGGCAGACCCCTCCAGATCCCGCCAGCAACCTTT 300
Db :::::::::::::::::::: |||||
234 ValThrProIleProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY 301 GCCCACTTACAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db :::::::::::::::::::: |||||
254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY 361 GTCCTGTGTTCTCGAGCTGGCGGAGACAGATCGCCCTCTCTGAAGGCATCCACT 420
Db :::::::::::::::::::: |||||

Db 274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY 421 ATCAGATCATCTGCTAGACAGCAGCGGTACACACAGCAGAGTGTATCACC 480
Db :::::::::::::::::::: |||||
294 IleGluIleMetLeuLeuGlnThrAlaArgTyrAsnHisGluThrGluCysIleThr 313
QY 481 TTCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTCAGGCGCTGCAGGTGGAG 539
Db :::::::::::::::::::: |||||
314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY 540 TTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATCGCGCGCTGGCGCTGGACGAGCT 599
Db :::::::::::::::::::: |||||
334 PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla 353
QY 600 GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGCCGACCGGCCCAACGTCAGGAG 659
Db :::::::::::::::::::: |||||
354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY 660 CCGGGCGCGCTGAGGCGCTTGCGAGCAGCCCTACGTGGAGCGCTGTGTCTCTACACGCGC 719
Db :::::::::::::::::::: |||||
374 ProSerArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY 720 ATCAAGAGCGCGCAGCAGCAGCTGCGC 746
Db :::::::::::::::::::: |||||
394 IleLysArgProGlnAspGlnLeuArg 402

RESULT 15

US-08-342-411A-4
Sequence 4, Application US/08342411A
Patent No. 5639616
GENERAL INFORMATION:
APPLICANT: LIAO, Shutsung
APPLICANT: SONG, Ching
TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR:
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/342,411A
FILING DATE: 18-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KITCHELL, BARBARA S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD154
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-342-411A-4

Alignment Scores:
Pred. No.: 4.12e-92 Length: 443
Score: 1094.00 Matches: 224
Percent Similarity: 92.37% Conservative: 6
Best Local Similarity: 89.96% Mismatches: 13
Query Match: 79.62% Indels: 7

DB:		1	Gaps:	1
US-10-712-629B-2 (1-746) x US-08-342-411A-4 (1-443)				
QY	1	AAGATTTCGAAACAGCAGCAGCAGGAGTCAAGTCCACAGTCGCAGCTCACCTGTGGCGCG	60	
Db	157	LyslleGlnLysGlnGlnGln-----GlnProProProThr	170	
QY	61	CAGGCAGCAGCAGCTCAGCCTCTGGGCTTGCGGTTCCTGGTGTGAATCTGAGCAGGC	120	
Db	171	GluProAlaSerGlySerSerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer	190	
QY	121	AGCCAGGGCTCGGGAGGCGAGGTGTCCAGCTAACAGCGGCTCAGAAGTAATAATGATC	180	
Db	191	SerGlnGlySerGlyGluGlyGluGlyLeuThrAlaAlaGlnGluLeuMetIle	210	
QY	181	CAGCAGTTGGTGGCGGCCAACCTGCAGTGTCAAACAACGCTCTCTCCGACCAGGCCAAA	240	
Db	211	GlnGlnLeuValAlaValGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys	230	
QY	241	GTCAGCCCTGGCCCTGGGGCGCAGACCCCCAGTCCCAGATGCCCGCCAGCAACGCTTT	300	
Db	231	ValThrProTrpProLeuGlyAlaAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe	250	
QY	301	GCCCATTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAA	360	
Db	251	AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln	270	
QY	361	GTGCTGGTTTCTGCAGCTGGCGCGGAGGACCAGATCGCCCTCTCGAAGCATCCACT	420	
Db	271	ValProGlyPheLeuGlnLeuGlyArgGluaspGlnIleAlaLeuLysAlaSerThr	290	
QY	421	ATCGAGATCATGCTCTAGACACGCGAGCGCTCACACCACGACAGACAGTGTATCAC	480	
Db	291	IleGluIleMetLeuLeuGluThrAlaArgArgTyraAsnHisGluThrGluCysilethr	310	
QY	481	TTCCTTGAG-GACTTCACCTACAGCAAGCAGACTTCCACCCTGCAGGCTGCAGGTGGAG	539	
Db	311	PheLeuLysAspPheThrTyrseryAspappPheHisArgAlaGlyLeuGlnValGlu	330	
QY	540	TTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGCTGGACACGCT	599	
Db	331	PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAla	350	
QY	600	GAGTACGCGCTGCTCATCGCCATCAACATCTTCTCGGCGGACCGGCCAACGTGCAGG	659	
Db	351	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	370	
QY	660	CCGGCCCGCTGGAGCGGTTCGACAGACGCCCTACGTGGAGCGCTGCTGTCTACACGCG	719	
Db	371	ProSerArgValGluAlaLeuGlnGlnProTyrrValGluAlaLeuLeuSertyrThrArg	390	
QY	720	ATCAAGAGCGCCAGGACCATGCTGCGC	746	
Db	391	IleLysArgProGlnAspGlnLeuArg	399	

Search completed: November 25, 2005, 22:33:18
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:23:04 ; Search time 150 Seconds
(without alignments)
4156.009 Million cell updates/sec

Title: US-10-712-629B-2

Perfect score: 1374

Sequence: 1 aagattcggaacagcagca.....ggccgcaggaccagctgcgc 746

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DSV=xlp
-Q=/cgn2/US70.spool_p/US10712629/runat 25112005 143038 11020/app_query.fasta_1.903
-DB=Published Applications AA Main -QMT=faetan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=biosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPEXT=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10712629 @CGN 1 1 307 @runat 25112005 143038 11020 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA Main.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	89.6	461	4	US-10-013-823-3
2	1231	89.6	461	4	US-10-418-007-4
3	1231	89.6	461	5	US-10-899-458-12
4	1231	89.6	461	5	US-10-500-912-4
5	1215.5	88.5	460	4	US-10-327-813-4
6	1215.5	88.5	460	4	US-10-329-668-4
7	1215.5	88.5	460	4	US-10-717-049-4
8	1215.5	88.5	460	5	US-10-509-197-4
9	1162	84.6	297	3	US-09-925-297-635
10	1101	80.1	446	3	US-09-909-446-2
11	1101	80.1	446	3	US-09-909-325-2

ALIGNMENTS

RESULT 1

US-10-013-823-3
; Sequence 3, Application US/10013823
; Publication No. US20020116731A1
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; APPLICANT: Zhang, Qin
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING RETINOID X
; TITLE OF INVENTION: RECEPTOR INTERACTING PROTEIN GENE DISRUPTIONS
; FILE REFERENCE: R-684
; CURRENT APPLICATION NUMBER: US/10/013,823
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/254,801
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/309,404
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-823-3

Alignment Scores:
Pred. No.: 7.69e-88 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 99.59% Indels: 1
DB: 4 Gaps: 0

12	1101	80.1	446	3	US-09-909-326-2	Sequence 2, Appli
13	1098.5	79.9	446	4	US-10-013-823-2	Sequence 2, Appli
14	1098.5	79.9	446	5	US-10-899-458-10	Sequence 10, Appl
15	1094	79.6	443	5	US-10-899-458-8	Sequence 8, Appli
16	1005	73.1	263	4	US-10-418-007-8	Sequence 8, Appli
17	777	56.6	264	4	US-10-043-487-254	Sequence 254, App
18	719.5	52.4	445	5	US-10-899-458-4	Sequence 4, Appli
19	718.5	52.3	445	5	US-10-899-458-2	Sequence 2, Appli
20	707.5	51.5	377	4	US-10-236-417-4	Sequence 4, Appli
21	705.5	51.3	553	4	US-10-338-411-25	Sequence 25, Appl
22	705.5	51.3	553	4	US-10-389-640-25	Sequence 25, Appl
23	704.5	51.3	451	4	US-10-236-417-8	Sequence 8, Appli
24	702.5	51.1	447	4	US-10-327-813-2	Sequence 2, Appli
25	702.5	51.1	447	4	US-10-329-668-2	Sequence 2, Appli
26	702.5	51.1	447	4	US-10-418-007-2	Sequence 2, Appli
27	702.5	51.1	447	4	US-10-429-160-12	Sequence 12, Appl
28	702.5	51.1	447	4	US-10-236-417-2	Sequence 2, Appli
29	702.5	51.1	447	4	US-10-236-417-6	Sequence 6, Appli
30	702.5	51.1	447	4	US-10-717-049-2	Sequence 2, Appli
31	702.5	51.1	447	5	US-10-899-458-6	Sequence 6, Appli
32	702.5	51.1	447	5	US-10-921-023-2	Sequence 2, Appli
33	702.5	51.1	447	5	US-10-500-912-2	Sequence 2, Appli
34	702.5	51.1	447	5	US-10-509-197-2	Sequence 2, Appli
35	692	50.4	433	4	US-10-278-945-2	Sequence 2, Appli
36	691	50.3	296	5	US-10-921-023-39	Sequence 39, Appl
37	680	49.5	191	4	US-10-104-047-2537	Sequence 2537, Ap
38	679	49.4	220	4	US-10-418-007-6	Sequence 6, Appli
39	660.5	48.1	511	5	US-10-921-023-4	Sequence 4, Appli
40	607.5	44.2	409	5	US-10-921-023-6	Sequence 6, Appli
41	504	36.7	136	5	US-10-921-023-27	Sequence 27, Appl
42	462	33.6	200	5	US-10-921-023-28	Sequence 28, Appl
43	400.5	29.1	316	4	US-10-468-199-74	Sequence 74, Appl
44	383	27.9	66	3	US-09-864-761-48359	Sequence 48359, A
45	367	26.7	167	4	US-10-424-599-177842	Sequence 177842,

```

US-10-712-629B-2 (1-746) x US-10-013-823-3 (1-461)
QY 1 AAGATTCGGAACACAGCAGCAGCAGCAGTCCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LyslleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCAGCCTCTGGGCTTGGGCTTCCCTGTGGATCTGAGGAGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAGCGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CACAGTTGTGGGGCCCAACTGCAGTGCAACAAACGCTCTCTCCGACCAAGCCCAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCTTGGGCTTGGGCGGAGACCCCAAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCACCTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTGCTTTCCTGCAGCTGGGCGGAGGACCAAGTCCGCTCTCTGAAGGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGACAGCAGCGCGCTACAAACAGCAGTCCGCTTCCGACCAAGC 480
Db 369 GlutyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCTGGAGGCGTTGACAGCGCTTCCAGCGGCGCTGCTGCTGCTACACGCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTrpValGluAlaLeuLeuSerThrArg 408
QY 720 ATCAAGAGCGCGCAGGACCAAGTCCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 2
US-10-418-007-4
; Sequence 4, Application US/10418007
; Publication No. US20040018560A1
; GENERAL INFORMATION:
; APPLICANT: Bledsoe, Randy
; APPLICANT: Miller, Ann B.
; APPLICANT: Moore, John
; APPLICANT: Moore, Linda
; APPLICANT: Williams, Shawn P.
; APPLICANT: Wisely, George B.
; TITLE OF INVENTION: CRYSTALLIZED LXR POLYPEPTIDE IN COMPLEX WITH A LIGAND AND SCREENING METHOD
; FILE REFERENCE: PU04691
; CURRENT APPLICATION NUMBER: US/10/418.007
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4

```

```

; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-418-007-4

Alignment Scores:
Pred. No.: 7.69e-88 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 4 Gaps: 0

US-10-712-629B-2 (1-746) x US-10-418-007-4 (1-461)
QY 1 AAGATTCGGAACACAGCAGCAGCAGCAGTCCACAGTCACAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 LyslleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGGCAGCAGCAGCAGCCTCTGGGCTTGGGCTTCCCTGTGGATCTGAGGAGGC 120
Db 189 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGGCTCCGGGAGCGAGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db 209 SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CACAGTTGTGGGGCCCAACTGCAGTGCAACAAACGCTCTCTCTCCGACCAAGCCCAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCACGCTTGGGCTTGGGCGGAGACCCCAAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCACCTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTGCTGCTTTCCTGCAGCTGGGCGGAGGACCAAGTCCGCTCTCTGAAGGATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCAGATCATGCTGTAGACAGCAGCGCGCTACAAACAGCAGTCCGCTTCCGACCAAGC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 328
QY 481 TTCTTGAG-CACATTACCTACAGCAGGACGACTTCCACCGTGCAGGCGCTGCGAGTGGAG 539
Db 329 PheLeuLysAspPheThrTyrSerLysAspPhePheHisArgAlaGlyLeuGlnValGlu 348
QY 540 TTCATCAACCCCATCTTCGAGTTCCTCGGGCGCATCGCGGCTTGGGCTTGGGCGCTTGGAG 599
Db 349 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 368
QY 600 GAGTACGCTGCTCATCGCCATCAACATCTTCTCGGCGCAGCCGCGCCCAAGCTGCAGGAG 659
Db 369 GlutyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 388
QY 660 CCGGGCGCTGGAGGCGTTGACAGCGCTTCCAGCGGCGCTGCTGCTGCTACACGCGC 719
Db 389 ProGlyArgValGluAlaLeuGlnGlnProTrpValGluAlaLeuLeuSerThrArg 408
QY 720 ATCAAGAGCGCGCAGGACCAAGTCCGC 746
Db 409 IleLysArgProGlnAspGlnLeuArg 417

RESULT 3
US-10-899-458-12
; Sequence 12, Application US/10899458
; Publication No. US2005008011A1
; GENERAL INFORMATION:
; APPLICANT: Bayne, Christopher D.

```

```
; APPLICANT: Johnson, Alan T.
; APPLICANT: Lu, Shao-po
; APPLICANT: Mohan, Raju
; APPLICANT: Nyman, Michael E.
; APPLICANT: Schweiger, Edwin J.
; APPLICANT: Stevens, William C. Jr.
; APPLICANT: Wang, Haixia
; APPLICANT: Xie, Yinong
; TITLE OF INVENTION: MODULATORS OF LXR
; FILE REFERENCE: 980049.411C1
; CURRENT APPLICATION NUMBER: US/10/899,458
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 10/327,813
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-899-458-12

Alignment Scores:
Pred. No.: 7,69e-88 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 5 Gaps: 0

US-10-712-629B-2 (1-746) x US-10-899-458-12 (1-461)

QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGGCTCAAGAACTAAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCCAGCTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTCCAGGAGCTGGCCCATCTCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCCTGGTTCTCCTCAGCTGGCGGGAGGACCAAGATCCGCTCTCTGAAGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATCTCTGCTAGACAGCAGGCAGGCCCTCAACACGACGACAGAGTGATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyAsnHisGluThrGluCysIleThr 328

; APPLICANT: Johnson, Alan T.
; APPLICANT: Lu, Shao-po
; APPLICANT: Mohan, Raju
; APPLICANT: Nyman, Michael E.
; APPLICANT: Schweiger, Edwin J.
; APPLICANT: Stevens, William C. Jr.
; APPLICANT: Wang, Haixia
; APPLICANT: Xie, Yinong
; TITLE OF INVENTION: MODULATORS OF LXR
; FILE REFERENCE: 980049.411C1
; CURRENT APPLICATION NUMBER: US/10/899,458
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 10/327,813
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-899-458-12

Alignment Scores:
Pred. No.: 7,69e-88 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 5 Gaps: 0

US-10-712-629B-2 (1-746) x US-10-500-912-4 (1-461)

QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGGCTCAAGAACTAAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCCAGCTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTCCAGGAGCTGGCCCATCTCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCCTGGTTCTCCTCAGCTGGCGGGAGGACCAAGATCCGCTCTCTGAAGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATCTCTGCTAGACAGCAGGCAGGCCCTCAACACGACGACAGAGTGATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyAsnHisGluThrGluCysIleThr 328

; APPLICANT: Johnson, Alan T.
; APPLICANT: Lu, Shao-po
; APPLICANT: Mohan, Raju
; APPLICANT: Nyman, Michael E.
; APPLICANT: Schweiger, Edwin J.
; APPLICANT: Stevens, William C. Jr.
; APPLICANT: Wang, Haixia
; APPLICANT: Xie, Yinong
; TITLE OF INVENTION: MODULATORS OF LXR
; FILE REFERENCE: 980049.411C1
; CURRENT APPLICATION NUMBER: US/10/899,458
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 10/327,813
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-899-458-12

Alignment Scores:
Pred. No.: 7,69e-88 Length: 461
Score: 1231.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 89.59% Indels: 1
DB: 5 Gaps: 0

US-10-712-629B-2 (1-746) x US-10-500-912-4 (1-461)

QY 1 AAGATTCGGAACAGCAGCAGCAGGAGTCACAGTCACAGTCGAGTCACCTGTGGGGCCG 60
Db 169 LysIleArgLysGlnGlnGlnGlnGlnSerGlnSerGlnSerProValGlyPro 188
QY 61 CAGGCGCAGCAGCAGCTCAGCTCTGGGCTGGGGCTTCCCTGGTGGATCTGAGGCAGGC 120
Db 189 GlnGlySerSerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 208
QY 121 AGCCAGGCTCCGGGAAGCGAGGTGTCCAGCTAACAGCGGCTCAAGAACTAAATGATC 180
Db 209 SerGlnGlySerGlyGlyGlyValGlnLeuThrAlaAlaGlnGluLeuMetIle 228
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 240
Db 229 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 248
QY 241 GTCCAGCTTGGTGGCGGCCCAACTGCAGTCAGTCAAAACGCTCTCTTCCGACACGCCCAA 300
Db 249 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 268
QY 301 GCCCACTTCCAGGAGCTGGCCCATCTCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 269 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 288
QY 361 GTCCCTGGTTCTCCTCAGCTGGCGGGAGGACCAAGATCCGCTCTCTGAAGCATCCACT 420
Db 289 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 308
QY 421 ATCGAGATCATCTCTGCTAGACAGCAGGCAGGCCCTCAACACGACGACAGAGTGATCACC 480
Db 309 IleGluIleMetLeuLeuGluThrAlaArgArgTyAsnHisGluThrGluCysIleThr 328
```


Qy	121	AGCCAGGGCTCGGGRAGGCGAGGTCCTCAGCTTAAACAGCGGCTCAAGAACTAATGATC	180
Db	208	SerGlnGlySerGlyGluGlyValGlnLeuThrAlaAlaGlnLeuMetIle	227
Qy	181	CAGCAGTTGGTGGCGGCCAACTGCAGTGCACAAACAGCTCCTCTCCGACCAGGCCAAA	240
Db	228	GlnGlnLeuValAlaAlaGlnLeuGlnCysAenLysArgSerPheSerAspGlnProLys	247
Qy	241	GTCAGCCCTGGCCCTGGCGGCAGACCCCGAGTCCCGAGATGCCCGCCAGCAACGCTTT	300
Db	248	ValThrProTyrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe	267
Qy	301	GCCCACTTCACGGAGCTGGCCATCATCTCAGTCCAGAGAGTCGTGGACTTCGCTAAGCAA	360
Db	268	AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln	287
Qy	361	GTGCTGTGTTCTGTCAGCTGGCGCGGAGACACAGATCGCCCTCCTGAAGGCATCCACT	420
Db	288	ValProGlyPheLeuGlnLeuGlyArgGlnAspGlnIleAlaLeuLeuLysAlaSerThr	307
Qy	421	ATCAGAGATCATGCTGTAGACACGCCAGGCGCTACAACACGACAGACAGAGTGTATCAC	480
Db	308	IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr	327
Qy	481	TTCCTTGAG-GACTTCACCTACAGCAAGACGACTTCACCGTCGAGGCGCTCGAGGTGAG	539
Db	328	PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu	347
Qy	540	TTCATCAACCCCATCTTCGAGTTCTCGGGGCCATGCGCGCGCTGGGGCTGGACGACCT	599
Db	348	PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyIleuLeuAspAla	367
Qy	600	GAGTACGCCCTGTCTATCGCCATCAACATCTTCTCGCGCCAGCCGCGCCCAACGTGCAGAG	659
Db	368	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	387
Qy	660	CCGGGCGCGTGGAGGCGTTTCAGCAGGCCCTACGTGGAGGCGCTGCTCTCTACACGGC	719
Db	388	ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg	407
Qy	720	ATCAAGAGCGCCAGGACCACTGGCG	746
Db	408	IleLysArgProGlnAspGlnLeuArg	416

Percent Similarity:	99.20%	Conservative:	0
Best Local Similarity:	99.20%	Mismatches:	1
Query Match:	88.46%	Indels:	2
DB:	4	Gaps:	1
US-10-712-629B-2 (1-746) x US-10-717-049-4 (1-460)			
Qy	1	AAGATTCCGGAACAGCAGCAGCAGGAGTCAAGTCAAGTCACAGTCGCGAGTCACCTGTGTGGGCGG	60
Db	169	LysileArgLys---GlnGlnGlnGlnSerGlnSerGlnSerGlnSerProValGlyPro	187
Qy	61	CAGGGCAGCAGCAGCTCAGCCTCTGGGCCTGGGCCTCCCTGGTGGATCTTGAGCAGGC	120
Db	188	GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly	207
Qy	121	AGCCAGGGCTCCGGGAAGCGAGGGTGTCCAGGCTAACAGCGGCTCAAGAACTAATGATC	180
Db	208	SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGlnLeuMetile	227
Qy	181	CAGCAGTTGTGGGGCCCAACTGCAGTGCMAACAACAGCTCCTTCTCCGACGAGCCCAA	240
Db	228	GlnGlnLeuValAlaAlaGlnLeuGlnCysAenlyAArgSerPheSerAspGlnProLys	247
Qy	241	GTCAGGCCCTGGCCCTGGGGCCAGACCCCAAGTCCCGAGATGCCCGCCAGCAACGCTTT	300
Db	248	ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe	267
Qy	301	GCCCACTTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCTGGAGCTTCGCTAAGCAA	360
Db	268	AlaHisPheThrGluLeuAlaIleIleSerValGlnIleValAspPheAlaLysGln	287
Qy	361	GTGCTTGTTCTCGAGCTGGGGCCGGAGGACCAGATCGCCCTCTCGAAGGCATCCACT	420
Db	288	ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr	307
Qy	421	ATCGAGATCATGCTGTAGACAGCCAGGGCGCTACAACACGACGACAGAGTGTATCACC	480
Db	308	IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr	327
Qy	481	TTCTTGAG-GACTTCACCTACAGCAAGCAGCACTTCACCGCTGCAGGCTCGCAGTGGAG	539
Db	328	PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu	347
Qy	540	TTCATCAACCCCATCTTCGAGTCTTCGCGGCCATCGCGCGCTGGGCTGGACGACGCT	599
Db	348	PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla	367
Qy	600	GAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGGCGCACCGGCCCAAGCTGAGAG	659
Db	368	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	387
Qy	660	CCGGGCGCTGGAGGGCTTCAGCAGCCCTACCTGGAGGGCGCTGCTCTCTACACGGC	719
Db	388	ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg	407
Qy	720	ATCAAGAGGCGCAGGACCAGCTGCGC	746
Db	408	IleLysArgProGlnAspGlnLeuArg	416
RESULT 8			
US-10-509-197-4			
; Sequence 4, Application US/10509197			
; Publication No. US20050171084A1			
; GENERAL INFORMATION:			
; APPLICANT: Burbidge, Stephen A.			
; APPLICANT: Cairns, William J.			
; APPLICANT: Irving, Elaine A.			
; APPLICANT: Parsons, Andrew A.			
; APPLICANT: Richardson, Jill C.			
; APPLICANT: Soden, Peter E.			
; APPLICANT: Vinson, Mary			
; APPLICANT: Watson, Mike A.			
; APPLICANT: Whitney, Karl D.			

```
; TITLE OF INVENTION: Methods of Treatment with LXR Modulators
; FILE REFERENCE: P51332
; CURRENT APPLICATION NUMBER: US/10/509,197
; CURRENT FILING DATE: 2004-09-24
; PRIOR APPLICATION NUMBER: PCT/US03/09225
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/368,424
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-509-197-4

Alignment Scores:
Pred. No.: 1,26e-86 Length: 460
Score: 1215.50 Matches: 247
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 88.46% Indels: 2
DB: 5 Gaps: 1

US-10-712-629B-2 (1-746) x US-10-509-197-4 (1-460)

Qy 1 AAGATTCGGAACAGCAGCAGGAGTACAGTCAGTCGCGAGTCACCTGTGGGGCCG 60
Db 169 Lysleahrgys--GlnGlnGlnSerGlnSerGlnSerProValGlyPro 187

Qy 61 CAGGGCAGCAGCAGCTCAGCTCGGCTCGGGCTTCCCTCGTGGTGTGAGTGTAGGCGGC 120
Db 188 GlnGlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGly 207

Qy 121 AGCCAGGCTCCGGGAGCGAGGTGTCAGTACAGCGCTCAAGAACTAAATGATC 180
Db 208 SerGlnGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMecile 227

Qy 181 CAGCAGTGTGTGGCGGCCCAACTGCAGTGCACAAACAAAGCTCTTCCGACCCAGCCCAA 240
Db 228 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 247

Qy 241 GTCAGCGCTTGGCGGCGAGACCCCGAGTCCCGAGATGCCGCCAGCAGCAGCTTT 300
Db 248 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 267

Qy 301 GCCCACTTCAGGAGTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTTAAGCAA 360
Db 268 AlahisPheThrGluLeuAlaIleIleSerValGlnGluLeuValAspPheAlaLysGln 287

Qy 361 GTCCTGTGTTCTCTGACGCTGGCGGAGGACCGATCGCCCTCTCTGAAGGATCCACT 420
Db 288 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 307

Qy 421 ATCAGATATGCTGTAGAGACGACCGGCTTACACACAGCAGACAGAGTGTATCACC 480
Db 308 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 327

Qy 481 TTCTTGAG-GACTTTCACCTACAGCAAGACGACTTCCACCGTCGAGCGCTCGAGTGAG 539
Db 328 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 347

Qy 540 TTCATCAACCCCATCTTTCGAGTTCTCGCGGCGCCATCGCGGCGCTGGCCCTGGACGCGT 599
Db 348 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 367

Qy 600 GAGTACGCCCTGTCTATCGCATCAACATCTTCTCGGCCGACCCGCCCAACGTCGAGAG 659
Db 368 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 387

Qy 660 CCGGGCCGCTGGAGGCGTTCACGACGACCCCTACGTCGAGGCGCTGCTCTACACGCGC 719
Db 388 ProGlyArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 407
```

```
Qy 720 ATCAAGAGCGCCGACAGGACCAGCTGCGC 746
Db 408 IleLysArgProGlnAspGlnLeuArg 416

RESULT 9
US-09-925-297-635
; Sequence 635, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (222)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (242)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (254)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (268)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (269)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (274)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (280)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (282)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (295)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (296)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-297-635

Alignment Scores:
Pred. No.: 1,83e-82 Length: 297
Score: 1162.00 Matches: 242
Percent Similarity: 97.19% Conservative: 0
Best Local Similarity: 97.19% Mismatches: 6
Query Match: 84.57% Indels: 3
DB: 3 Gaps: 0

US-10-712-629B-2 (1-746) x US-09-925-297-635 (1-297)

Qy 3 GATTCGGAACAGCAGCAGGAGTACAGTCAGTCAGTCACCTGTGGGGCCGCA 62
Db 8 AspSerGluThrThrAlaAlaGlyValThrValThrValAlaValThrCysGlyAlaAla 27

Qy 63 GGCACAGCAGCAGCTCAGCCCTTCGGGCTTCCTCGGTGGATCTGAGGCGCAG 122
```

```
Db      28  -GlySerSerSerAlaSerGlyProGlyAlaSerProGlyGlySerGluAlaGlySe 47
QY      123  CCAGGGCTCCGGGAGCGGAGGTGTCAGCTAACAGCGGCTCAAGAACTAATGATCCA 182
Db      47  rGInGlySerGlyGluGlyGluGlyValGlnLeuThrAlaAlaGlnGluLeuMetIleGI 67
QY      183  GCATTGGTGGCGGCCCAACTGCAGTGCACAAACGCTCTTCTCCGACCCAGCCCAAGT 242
Db      67  nGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLysVa 87
QY      243  CAGCCCTGGCCCTCGGGCGCAGACCCCGAGTCCCGAGATCCCGCCGCGCAAGCGTTTGC 302
Db      87  lThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPheAl 107
QY      303  CCACTTCACGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAAGT 362
Db      107  aHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGlnVa 127
QY      363  GCTGTTTCTCCAGCTGGCGCGGAGGACCAAGATCGCCCTCTGAAAGCATCCACTAT 422
Db      127  lProGlyPheLeuGlnGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThrIl 147
QY      423  CGAGATCATGCTCTAGAGACAGCCAGCGGCGCTACAACCCAGGACAGAGTGATACCTT 482
Db      147  eGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThrPh 167
QY      483  CTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTGACGGCTCGAGTGAGTT 541
Db      167  eLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGluPh 187
QY      542  CATCAACCCCATCTTCGAGTTCTCGGGGCGCATCGCGCGCTGGCGCTGGACGCTGA 601
Db      187  eIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAlaGI 207
QY      602  GTACGGCTGCTCATCGCCATCAACATCTTCTCGGCGGACCGGCCCAAGCTCGAGGCC 661
Db      207  uTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArg***AsnValGlnGluPr 227
QY      662  GSGCGCGTGGAGCGCTTCGACGAGCCCTACGT-GGAGGCGTCTGCTACACGGCA 720
Db      227  oGlyArgValGluAlaLeuGlnGlnProTyrValGlyGlyAla***ValLeuHisAlaH 247
QY      721  TCAAGAGCGCGCAGGACCAAGCTCGC 745
Db      247  sGlnGluAlaAlaGlyPro***Ala 255
```

RESULT 10

US-09-909-446-2

; Sequence 2, Application US/09909446

; Patent No. US20020052489A1

; GENERAL INFORMATION:

; APPLICANT: ENMARK, EVA

; TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING

; TO THE NUCLEAR RECEPTOR FAMILY

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Banner & Witcoff

; STREET: 1001 G Street, NW

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/909,446

; FILING DATE: 19-Jul-2001

; CLASSIFICATION: <Unknown>

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/776,844
; FILING DATE: <Unknown>
; APPLICATION NUMBER: UK 9413536.2
; FILING DATE: 16-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 00487.04029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-446-2
```

```
Alignment Scores:
Pred. No.: 1.18e-77 Length: 446
Score: 1101.00 Matches: 225
Percent Similarity: 92.77% Conservative: 6
Best Local Similarity: 90.36% Mismatches: 12
Query Match: 80.13% Indels: 7
DB: Gaps: 1
```

US-10-712-629B-2 (1-746) x US-09-909-446-2 (1-446)

```
QY      1  AGATTCCGAACACAGCAGCAGCAGGAGTCACATGCAGTCGAGTCACCTGTGGGGCCG 60
Db      160  LysIleGlnLysGlnGlnGln-----GlnProProProThr 173
QY      61  CAGGGCAGCAGCAGCTCGGCTCTGGGCTTCCCTGGGTCTCGATCTGAGGCGAGGC 120
Db      174  GluProAlaSerGlySerSerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY      121  AGCCAGGGCTCCGGGAAGCGAGGTGTCCAGCTAACACGCGCTCAAGAACTAATGATC 180
Db      194  SerGlnGlySerGlyGlyGlyGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 213
QY      181  CAGCAGTTGCTGGCGGCCCAACTGCAGTGCACAAACGCTCTCTCCGACCGCCCAAA 240
Db      214  GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY      241  GTCAGCCCTCGCCCTGGCGCAGACCCAGTCCCGAGATGCCCGCCAGCAACGCTTT 300
Db      234  ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 253
QY      301  GCCCACTTCACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db      254  AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY      361  GTGCCCTGTTTCTCGAGCTGGCGGAGGAGGACCATCGCCCTCTCTGAAGGCATCACT 420
Db      274  ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY      421  ATCGAGATCATGCTGTAGACAGCGCGCTACAACCCAGACAGACAGAGTGTATCACC 480
Db      294  IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY      481  TTCCTTGAG-GACTTCACCTACAGCAAGGACGACTTCCACCGTCGAGCGCTGCAAGTGGAG 539
Db      314  PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 333
QY      540  TTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGGCTGGGCTGGACGCGCT 599
Db      334  PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAla 353
```


Db	334	PheIleAsnProIlePheGluPheSerArgAlaMetArgLeuGlyLeuAspAspAla	353
Qy	600	GAGTACGCGCTCATCGCCATCAACATCTTCTCGGCGCACCGGCCCAACGTGCAGGAG	659
Db	354	GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu	373
Qy	660	CCGGGCGCGTGGAGGGGTTTCAGCAGCGCTTACGTGGAGGGCGCTCTGTCTACACGCGC	719
Db	374	ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg	393
Qy	720	ATCAAGAGGCGCGAGGACCACTGCGC	746
Db	394	IleLysArgProGlnAspGlnLeuArg	402

RESULT 13

US-10-013-823-2

; Sequence 2, Application US/10013823

; Publication No. US20020116731A1

; GENERAL INFORMATION:

; APPLICANT: Guenther, Catherine

; APPLICANT: Phillips, Russell

; APPLICANT: Allen, Keith D.

; APPLICANT: Zhang, Qin

; APPLICANT: Baribault, Helene

1 TITLE OF INVENTION: TRANSGENIC MICE CONTAINING RETINOID X
 2
 3
 4
 5 TITLE OF INVENTION: RECEPTOR INTERACTING PROTEIN GENE DISRUPTIONS
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520

; PRIOR FILING DATE: 2001-07-31
 ; NUMBER OF SEQ ID NOS: 5

```
; SOFTWARE: FASTSEQ LUI WINDOWS VERSION 4.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-013-823-2

Alignment Scores:
Pred. No.:          Length:      446
Score:             Matches:     224
Percent Similarity: 93.57%       Conservative:   9
Best Local Similarity: 89.96%    Mismatches:    15
Query Match:        Indels:      2
DB:                 Gaps:         1

US-10-712-629B-2 (1-746) x US-10-013-823-2 (1-446)

Qy      1 AAGATTCTGGAAACAGCAGCAGGAGTCACAGTCCAGTCACTGTGGGGCGG 60
           :::|::|||:::::::::::||::|::|
Db      155 GlNlEArgLyStYsArgIleGlNLys---GIgNGlnGInGInProPProSer 173

Qy      61 CAGGGCAGCAGCACACTCAGCCTCTGGGCCTGCCCTTGTTGGATCTGAGCAGCG 120
           ::::|::|::|::|::|::|::|::|
Db      174 GluPrOAlAIAIsrSrSrgLYArgrOlAlAsErProGLYThrSeRglUlAser 193

Qy      121 AGCCAGGGCTCCGGGGGAAGCGAGGGTGTCCAGCTAACAGCGGCTCAAAGAATAATGATC 180
           ||::|::|::|::|::|::|::|::|
Db      194 SerGIgNglySerGIyGUgLUyGuLIyleGlnLeuthrAlAIAIgLnGluLeuMetile 213

Qy      181 CAGCAGTTGGTGGCGGCCCAACTCCAGTGCAACAACAGCTCTCTTC CGACGAGCCCCAAA 240
           |::|::|::|::|::|::|::|::|
Db      214 GIgNGlnLeuValAIAlAgInLeuGIcNsASnLysArgSerPhESerASPgnProlys 233

Qy      241 GTACGCCCTTGGCCCCTGGGGCGAGACC CCAGTCCCGAGATGCCCGCAGCAACGCTTT 300
           |::|::|::|::|::|::|::|::|
Db      234 ValthrProTrpProLeUGiyAlAaspproGINserArgAspAlAargglnglnmargphe 253

Qy      301 GCCCACATTACGGAGCTGGGCCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTTAAGCAA 360
```

```
Db      254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY      361 GTGCGCTGGTTTCTCGCAGCTGGGCGGAGGACAGATCGCCCTCTCGAAGGCATCCACT 420
Db      274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY      421 ATCGAGATCATGCTGTAGACAGCCAGCGCGTCAACACAGAGACAGAGTGTATCACCC 480
Db      294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY      481 TTCTTTGAG-GACTTCACTACAGCAGGACGACTTCCACCGTCGAGCGCTCGAGGTGGAG 539
Db      314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY      540 TTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGCGCTGGACGCGCT 599
Db      334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 353
QY      600 GAGTAGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
Db      354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY      660 CCGGGCCGCGTGAGCGCTTGAGCAGCGCCCTACGTGAGGCGCTGCTGTCTACACGCGC 719
Db      374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY      720 ATCAAGAGCGCGCAGGACCGAGCTGCGC 746
Db      394 IleLysArgProGlnAspGlnLeuArg 402
```

RESULT 14

```
US-10-899-458-10
; Sequence 10, Application US/10899458
; Publication No. US20050080111A1
; GENERAL INFORMATION:
; APPLICANT: Bayne, Christopher D.
; APPLICANT: Johnson, Alan T.
; APPLICANT: Lu, Shao-po
; APPLICANT: Mohan, Raju
; APPLICANT: Nyman, Michael C.
; APPLICANT: Schweiger, Edwin J.
; APPLICANT: Stevens, William C. Jr.
; APPLICANT: Wang, Haixia
; APPLICANT: Xie, Yinong
; TITLE OF INVENTION: MODULATORS OF LXR
; FILE REFERENCE: 980049.411C1
; CURRENT APPLICATION NUMBER: US/10/899,458
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 10/327,813
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-899-458-10
```

```
Alignment Scores:
Pred. No.: 1.85e-77 Length: 446
Score: 1098.50 Matches: 224
Percent Similarity: 93.57% Conservative: 9
Best Local Similarity: 89.96% Mismatches: 15
Query Match: 79.95% Indels: 2
DB: 5 Gaps: 1
```

US-10-712-629b-2 (1-746) x US-10-899-458-10 (1-446)

```
QY      1 AAGATTCGGAACAGCAGCAGCAGGAGTCACTGTCAGTCGAGTCACCTGTGGGCGC 60
      :::::::::::::::::::: |||:: ||| ||| |||
```

```
Db      155 GlnIleArgLysLysArgIleGlnLys---GlnGlnGlnGlnGlnProProProSer 173
QY      61 CAGGCGCAGCAGCAGCTCAGCCTCTGGGCTTGGGGCTTCCCTTGGTGGATCTGAGCGAGC 120
      :::::::::::::::::::: |||
Db      174 GluProAlaAlaSerSerSerGlyArgProAlaAlaSerProGlyThrSerGluAlaSer 193
QY      121 AGCCAGGCTCCGGGAGGCGAGGTGTCTCAGCTTAACAGCGCTCAAGAACTTAATGATC 180
Db      194 SerGlnGlySerGlyGluGlyGluGlyIleGlnLeuThrAlaAlaGlnLeuMetIle 213
QY      181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGAACAAACGCTCTCTTCCGACACGAGCCAAA 240
Db      214 GlnGlnLeuValAlaAlaGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 233
QY      241 GTCAGCGCTGCGCCCTGGCGCAGACCCCAAGTCCCGAGATGCCCGCAGCAACCGCTTT 300
Db      234 ValThrProTyrProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnArgPhe 253
QY      301 GCCCACTTCACGAGCTGCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db      254 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 273
QY      361 GTGCGCTGGTTTCTCGCAGCTGGCGGAGGAGCAGATCGCCCTCTCGAAGGCATCCACT 420
Db      274 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 293
QY      421 ATCGAGATCATGCTGTAGACAGCAGCGCGCTCAACACAGAGACAGAGTGTATCACCC 480
Db      294 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 313
QY      481 TTCTTTGAG-GACTTCACTACAGCAGGACGACTTCCACCGTCGAGCGCTGCGAGTGGAG 539
Db      314 PheLeuLysAspPheThrTyrSerLysAspPheHisArgAlaGlyLeuGlnValGlu 333
QY      540 TTTCATCAACCCCATCTTCGAGTTCTCGCGGGCCATCGCGCGCTGGCGCTGGACGAGC 599
Db      334 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 353
QY      600 GAGTAGCGCTGCTCATCGCCATCAACATCTTCTCGCGCGACCGCGCCCAACGTCGAGGAG 659
Db      354 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 373
QY      660 CCGGGCCGCGTGAGCGCTTGAGCAGCGCCCTACGTGAGGCGCTGCTGTCTACACGCGC 719
Db      374 ProSerArgValGluAlaLeuGlnGlnProTyrValGluAlaLeuLeuSerTyrThrArg 393
QY      720 ATCAAGAGCGCGCAGGACCGAGCTGCGC 746
Db      394 IleLysArgProGlnAspGlnLeuArg 402
```

RESULT 15

```
US-10-899-458-8
; Sequence 8, Application US/10899458
; Publication No. US20050080111A1
; GENERAL INFORMATION:
; APPLICANT: Bayne, Christopher D.
; APPLICANT: Johnson, Alan T.
; APPLICANT: Lu, Shao-po
; APPLICANT: Mohan, Raju
; APPLICANT: Nyman, Michael C.
; APPLICANT: Schweiger, Edwin J.
; APPLICANT: Stevens, William C. Jr.
; APPLICANT: Wang, Haixia
; APPLICANT: Xie, Yinong
; TITLE OF INVENTION: MODULATORS OF LXR
; FILE REFERENCE: 980049.411C1
; CURRENT APPLICATION NUMBER: US/10/899,458
; CURRENT FILING DATE: 2004-07-24
; PRIOR APPLICATION NUMBER: US 10/327,813
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,707
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 27
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-899-458-8
```

```
Alignment Scores:
Pred. No.: 4.17e-77 Length: 443
Score: 1094.00 Matches: 224
Percent Similarity: 92.37% Conservative: 6
Best Local Similarity: 89.96% Mismatches: 13
Query Match: 79.62% Indels: 7
DB: 5 Gaps: 1
```

US-10-712-629B-2 (1-746) x US-10-899-458-8 (1-443)

```
QY 1 AGATTTCGGAACAGCAGCAGCAGGAGTCACTCAGTCGAGTCACCTGTGGGGCCG 60
Db |||||:|||||||
157 LysIleGlnLysGlnGlnGln-----GlnProProProThr 170
QY 61 CAGGGCAGCAGCAGCTCAGCCTCGGCCCTGGGGCTTCCCTGGTGATCTGAGGAGGC 120
Db |||||:|||||||
171 GluProAlaSerGlySerAlaArgProAlaAlaSerProGlyThrSerGluAlaSer 190
QY 121 AGCCAGGGCTCCGGGAAGCGAGGGTGTCCAGCTAACAGCGCTCAAGAACTAATGATC 180
Db |||||:|||||||
191 SerGlnGlySerGlyGluGlyIleGlnLeuThrAlaAlaGlnGluLeuMetIle 210
QY 181 CAGCAGTTGGTGGCGGCCCAACTGCAGTGCACAAACAGTCTTCTCCGACCAGCCCAA 240
Db |||||:|||||||
211 GlnGlnLeuValAlaValGlnLeuGlnCysAsnLysArgSerPheSerAspGlnProLys 230
QY 241 GTCACGCCCTGGCCCTGGCGGAGAGCCCGAGTCCGAGATCCCGCAGCAACGCTTT 300
Db |||||:|||||||
231 ValThrProTrpProLeuGlyAlaAspProGlnSerArgAspAlaArgGlnGlnArgPhe 250
QY 301 GCCCACTTCACGAGGTGGGCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db |||||:|||||||
251 AlaHisPheThrGluLeuAlaIleIleSerValGlnGluIleValAspPheAlaLysGln 270
QY 361 GTGCTGGTTCTGTCAGTGGCGGAGAGCACAGATCGCCCTCTGAAGGCATCCACT 420
Db |||||:|||||||
271 ValProGlyPheLeuGlnLeuGlyArgGluAspGlnIleAlaLeuLeuLysAlaSerThr 290
QY 421 ATCGAGATCATGCTGTAGACAGCAGCGCTACACCAACGACAGACAGAGTGTATCACC 480
Db |||||:|||||||
291 IleGluIleMetLeuLeuGluThrAlaArgArgTyrAsnHisGluThrGluCysIleThr 310
QY 481 TTCCTTGAG-GACTTCACCTACAGCAAGGAGCACTTCCACCGTCGACGGCTGCAGGTGGAG 539
Db |||||:|||||||
311 PheLeuLysAspPheThrTyrSerLysAspAspPheHisArgAlaGlyLeuGlnValGlu 330
QY 540 TTCATCAACCCCATCTTCAGATTCTCGCGGCCCATCGCGGCTGGGCTGGACGACGCT 599
Db |||||:|||||||
331 PheIleAsnProIlePheGluPheSerArgAlaMetArgArgLeuGlyLeuAspAspAla 350
QY 600 GAGTAGCCCTGCTCATCGCCATCACATCTTCTCGCGCGAGCCGGCCCAACGTGCAGGAG 659
Db |||||:|||||||
351 GluTyrAlaLeuLeuIleAlaIleAsnIlePheSerAlaAspArgProAsnValGlnGlu 370
QY 660 CCGGGCGCGGTGGAGCGTTGCAGCAGCCCTACGTGGAGCGCTGTCTCTACACGCCGC 719
Db |||||:|||||||
371 ProSerArgValGluAlaLeuGlnProTyrValGluAlaLeuLeuSerTyrThrArg 390
QY 720 ATCAAGAGCCCGCAGGACCGAGCTGCGC 746
Db |||||:|||||||
391 IleLysArgProGlnAspGlnLeuArg 399
```

Search completed: November 25, 2005, 22:38:27
Job time : 161 secs

This Page Blank (Page 1000)

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (K9, P10)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2005, 22:17:17 ; Search time 4.5 Seconds

(without alignments)

1004.935 Million cell updates/sec

Title: US-10-712-629B-2

Perfect score: 1374

Sequence: 1 aagattcggaacacagcagca.....ggccgagaccagctggcg 746

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 35090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-MODEL=frame-n2p.model -DEV=xlp
-Q=/cn2_1/USFTO_spool_p/US10712629/runat_25112005_143038_11026/app_query.fasta_1.903
-DB=Published Applications AA New -QFMT=fastran -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10712629 @CN 1.1 @runat_25112005_143038_11026
-NCPUL6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases:

Published Applications AA New.*
1: /cn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
2: /cn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /cn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
7: /cn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.5	12.7	505	1	US-10-519-447-4
2	140.5	10.2	1467	1	US-10-821-234-1096
3	117.5	8.6	614	7	US-10-015-546A-20
4	117	8.5	1970	1	US-10-821-234-1641
5	116.5	8.5	1366	1	US-10-821-234-1431
6	113	8.2	735	7	US-11-053-100-44
7	112	8.2	1166	1	US-10-821-234-964
8	111	8.1	624	7	US-11-053-100-48
9	111	8.1	774	7	US-11-053-100-49
10	111	8.1	1560	7	US-11-059-982-1

11	108	7.9	1874	1	US-10-821-234-1182	Sequence 1182, Ap
12	107.5	7.8	1225	7	US-11-053-100-50	Sequence 50, Appl
13	105.5	7.7	775	7	US-11-053-100-51	Sequence 51, Appl
14	103.5	7.5	308	1	US-10-967-527A-30	Sequence 30, Appl
15	103.5	7.5	1366	1	US-10-821-234-1431	Sequence 1431, Ap
16	103	7.5	915	1	US-10-821-234-1514	Sequence 1514, Ap
17	102	7.3	138	1	US-10-667-295-176	Sequence 176, Appl
18	100.5	7.3	614	7	US-11-015-546A-20	Sequence 20, Appl
19	99.5	7.2	594	1	US-10-131-826A-10	Sequence 10, Appl
20	99.5	7.2	558	1	US-10-821-234-921	Sequence 921, Appl
21	99.5	7.2	1467	1	US-10-821-234-1096	Sequence 1096, Ap
22	98	7.1	757	7	US-11-053-100-47	Sequence 47, Appl
23	98	7.1	1076	1	US-10-131-826A-219	Sequence 219, Appl
24	97.5	7.1	888	1	US-10-131-826A-544	Sequence 544, Appl
25	96.5	7.0	766	1	US-10-821-234-1691	Sequence 1691, Ap
26	96	6.9	239	1	US-10-821-234-1186	Sequence 1186, Ap
27	96	6.9	1311	1	US-10-509-424-5	Sequence 5, Appl
28	96	6.9	1493	7	US-11-004-057-4	Sequence 4, Appl
29	95.5	7.0	736	7	US-11-053-100-45	Sequence 45, Appl
30	95.5	7.0	1186	7	US-11-053-100-46	Sequence 46, Appl
31	95	6.9	280	1	US-10-821-234-1300	Sequence 1300, Ap
32	95	6.9	1493	7	US-11-004-057-4	Sequence 4, Appl
33	95	6.9	1532	1	US-10-821-234-914	Sequence 914, Appl
34	94	6.8	575	1	US-10-131-826A-128	Sequence 128, Appl
35	93	6.8	405	1	US-10-821-234-1357	Sequence 1357, Ap
36	93	6.8	419	1	US-10-821-234-1556	Sequence 1556, Ap
37	93	6.7	594	1	US-10-131-826A-10	Sequence 10, Appl
38	93	6.7	1493	7	US-11-004-057-21	Sequence 21, Appl
39	92.5	6.7	479	1	US-10-821-234-871	Sequence 871, Appl
40	92.5	6.7	543	1	US-10-689-742-78	Sequence 78, Appl
41	91.5	6.7	1133	1	US-10-821-234-1219	Sequence 1219, Ap
42	91	6.6	1377	1	US-10-821-234-1070	Sequence 1070, Ap
43	91	6.6	1874	1	US-10-821-234-1182	Sequence 1182, Ap
44	90.5	6.6	503	7	US-11-013-247A-4	Sequence 4, Appl
45	90.5	6.6	1493	7	US-11-004-057-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1

US-10-519-447-4

; Sequence 4, Application US/10519447

; Publication No. US20050244829A1

; GENERAL INFORMATION:

; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.

; APPLICANT: Makoto OGINO

; APPLICANT: Hideki ENDOH

; TITLE OF INVENTION: METHOD FOR SCREENING AN AGENT FOR IMPROVING INSULIN RESISTANCE

; FILE REFERENCE: Q85576

; CURRENT APPLICATION NUMBER: US/10/519,447

; CURRENT FILING DATE: 2004-12-30

; PRIOR APPLICATION NUMBER: PCT/JP03/08367

; PRIOR FILING DATE: 2003-07-01

; PRIOR APPLICATION NUMBER: JP 2002-193814

; PRIOR FILING DATE: 2002-07-02

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 505

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-519-447-4

Alignment Scores:			
Pred. No.:	9.17e-07	Length:	505
Score:	174.50	Matches:	52
Percent Similarity:	47.13%	Conservative:	30
Best Local Similarity:	29.89%	Mismatches:	57
Query Match:	12.70%	Indels:	35
DB:	1	Gaps:	4

US-10-712-629B-2 (1-746) x US-10-519-447-4 (1-505)


```
QY 295 CCGTTTGGCCCACTTCACGAGCTG----- 318
Db 291 LysPheLeuHisIleThrProLeuGlnGlnSerLysGluValAlaIleArgIlePhe 310
QY 319 -----CCCATCATCTCAGTCCAGGAGATCGTGACTTCGCTAAGCAA 360
Db 311 GlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAlaLysSer 330
QY 361 GTGCGCTGGTTCTCGCAGCTGGCGGAGGACAGATCGCCCTCTCGAAGGCATCCACT 420
Db 331 IleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyVal 350
QY 421 ATCGAGATCATGCTGTAGACAGACGCGCTACAACACGAGACAGAGTGATATCAC 480
Db 351 HisGluIleIle-----Tyr-ThrMetLeuAlaSerLeuMetAs 363
QY 481 TTCTTGAGGACTTCACGCAAGGACGA-----CTTCCACC 519
Db 363 nLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSe 383
QY 520 GTGCAGG---CCTGCAGGTGGAGTTTCATCAACCCCATCTTCGAGTTCTCGCGGCCCATGC 576
Db 383 rLeuArgLysProPheGly-AppPheMetGluProLysPheGluPheAlaValLysPheA 403
QY 577 GCGCGCTGGCCCTGGACGACGCTAGTACGCCCTCTCATCGCCATCAACATCTTCTCGG 636
Db 403 snAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleLeuSerG 423
QY 637 CGACCGGCCCACTGTCAGGAGCCGCGCGCTGGAGCGGTTCGACGACCCCTACGTGG 696
Db 423 lYAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsnLeuLeuG 443
QY 697 AGGCGCTGCTCTACACGCGCATCAAGAGGCGCGCAG 734
Db 443 InAlaLeuGluLeuGlnLeuLysLeuAsnHisProGlu 455
```

RESULT 2

```
US-10-821-234-1096
; Sequence 1096, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; PRIOR FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1096
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1096
```

```
Alignment Scores:
Pred. No.: 0.000293 Length: 1467
Score: 140.50 Matches: 81
Percent Similarity: 36.16% Conservative: 17
Best Local Similarity: 29.89% Mismatches: 101
Query Match: 10.23% Indels: 74
DB: 1 Gaps: 16
```

US-10-712-629b-2 (1-746) x US-10-821-234-1096 (1-1467)

```
QY 49 CTTGTGGGGCGGCGGACGAGCTCAGCGCTCTGGGCTCTGGGCTTCCCTGTGGGA 108
Db 285 ProAlaGlyProLysGlyGluProGlySer-----ProGlyGluAsnGlyAlaPro 301
```

```
QY 109 TCTGAGGCGAGGAG-----CCAGGCTCCGGGAAGCGAGGGTGTCCAGCTAAC 158
Db 302 GlyGlnMetGlyProArgGlyLeuProGlyGluArgGlyArgProGlyAlaProGlyPro 321
QY 159 AGCGCTCAAGAACTAATGATCCAGCAGTTGGTGGC-----GGCCCAACTGC- 205
Db 322 AlaGlyAlaArgGlyAsnAspGlyAlaThrGlyAlaAlaGlyProProGlyProThrGly 341
QY 206 -----AGTGCAACAACGCTCTCTTCTCCGACC 232
Db 342 ProAlaGlyProProGlyPheProGlyAlaValAlaGlyAlaLysGlyGluAlaGlyProGln 361
QY 233 ACCCCAAAGTACGCCCTGGCCCTGGGGG-----CAGACCCCAAGTCCCGAGATGCC 286
Db 362 GlyProArgGlySerGluGlyProGlnGlyValArgGlyGluProGlyProGlyPro 381
QY 287 GCCACGAAGCTTTCGCCACTTCAGGAGCTGCCATCATCTCAGTCCAGAGAGATCGTGG 346
Db 382 AlaGlyAlaAlaGlyProAlaGlyAsnProGlyAlaAspGlyGln-ProGlyAlaLysG 401
QY 347 ACTTCGCTAAGCA-----AGTGCCTGGTTTCTGCAGCT---GGGCCGGG 388
Db 401 yAlaAsnGlyAlaProGlyIleAlaGlyAlaProGlyPheProGlyAlaArgGlyProSe 421
QY 389 AGGACCATCGCCCT-----CCTGAAGGCATCCACTATCGAGATCATGCTGCTAGAGA 442
Db 421 rGlyProGlnGlyProGlyProGlyProGly--ProLysGlyAsnSer-----Gly 437
QY 443 CAGCCAGGCGCTTACACCAACGAGAGAGTGTATCACCTTCTTAGGAGCTTCACCTACAG 502
Db 438 GluProGlyAlaProGlySerLysGlyAspThrGly-----Ala 450
QY 503 CAAGGACGACTTCCACCGTCGAGGCTGCAGGTGGAGTTTCATCAACCCCATCTTTCAGTT 562
Db 451 LysGlyGluProGlyProVal-----GlyValGlnGlyPro----- 462
QY 563 CTCGCGGCGCATCGCGCGCTGGGCT-----GGAGGAGCTGAGTACGCGCCT 610
Db 463 -----ProGlyProAlaGlyGluGluGlyLysArgGlyAlaArgGly 476
QY 611 GCTCATCGCCATCAACATCTTCTCGCGCGACCGGCCAACGTCGAGAGCGCGGCGCGCT 670
Db 477 GluProGlyProThrGlyLeuProGlyProProGlyGluArgGlyGlyProGlySerArg 496
QY 671 GGA-----GGCGTTGCAGCAGCCCTACGTGGAGGCGCTGTGTCTCTACAC 715
Db 497 GlyPheProGlyAlaAspGlyValAlaGlyPro---LysGlyProAla----- 511
QY 716 GGCATCAAGAGCGCGCAGGACGAGCT 742
Db 512 GlyGluArgGlySerProGlyProAla 520
```

RESULT 3

```
US-11-015-546A-20
; Sequence 20, Application US/11015546A
; Publication No. US20050250126A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; APPLICANT: APPELBY, MARK
; TITLE OF INVENTION: Ztnf13, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-24
; CURRENT APPLICATION NUMBER: US/11/015,546A
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,185
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 614
```


Db 490 ---GlnGlyAlaalaSerArgProAlaIleLeuTyAlaLeuLeuSerSerLeuLys 508
QY 270 CCAGTCCCGAGATGCCCGCAGCAACGCTTTGGCCACTTCACCGA----- 314
Db 509 AlaValProArgProArgSerArgCysLeuCysArgGlnHisArgProValGlnLeuCys 528
QY 315 GCTGGCCATCATCTCAGTCCAGGATCGTGGAC-----TTCGC 353
Db 529 AlaProHisArg-ThrCysArgGluAlaLeuAspValLeuAlaIleThrValAlaPheLe 548
QY 354 TAAGCAAGTCCCTGCTTCTCGCAGCTGGCCGGGAGGACCATCGCTCTCTGAAG-- 411
Db 548 uArgAenLeuProSerPheTrpGlnLeuProProGlnAspGlnArgLeuLeuGlnG 568
QY 412 -----GCATCCACTATCGAGATC-- 429
Db 568 yCysTrpGlyProLeuPheLeuLeuGlyLeuAlaGlnAspAlaValThrPheGluValAl 588
QY 430 -----ATGCTGTAGACAGCCAGCGCGCTACAA 458
Db 588 aGluAlaProValProSerIleLeuLysIleLeuLeuGlu-GluPro----- 604
QY 459 CCAGGAGACAGATGTATCACCTTCTTGAGGACTTCACCTACAGCAGGACGACTTCCA- 517
Db 605 -----SerSerSerGlyGly-----SerGlyGlnLeuProA 615
QY 518 -----CGTGCAGCGCTGCAGGTGGAGTTCATCAACCCCATCTTCGAGTTCTCG 566
Db 615 spArgProGlnProSerLeuAlaIleValGlnTrpLeuGlnCysCysLeuGlu----- 632
QY 567 CGGGCCATCGCGCGCTGGCGCTGGACGCTGAGTACGCCCTGCTCATCGCCATCAAC 626
Db 633 ---SerPheTrpSerLeuGluLeuSerProLysGluTyAlaCysLeuLysGlyThrIle 651
QY 627 ATCTTCTCGCGCAGCGGCCCAACGTGCAGAGCCCGCGCGCTGGAGCGTTTCAGCAG 686
Db 652 LeuPheAsnProAspValProGlyLeuGlnAlaAlaSerHisIleGlyHisLeuGlnGln 671
QY 687 ---CCCTACGTGGAGCGCTGC-----TGTCCTACACGCGCATCAAG 725
Db 672 GluAlaHis-TrpValLeuCysGluValLeuGluProTrpCysPro-----AlaAlaG 689
QY 726 AGCGCCGAGGACC 738
Db 689 nGlyArgLeuThr 693

RESULT 7
US-10-821-234-964
; Sequence 964, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presenclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 964
; LENGTH: 1166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-964

Alignment Scores:
Pred. No.: 0.0378 Length: 1166
Score: 112.00 Matches: 46
Percent Similarity: 35.12% Conservative: 13

Best Local Similarity: 27.38% Mismatches: 49
Query Match: 8.15% Indels: 60
DB: 1 Gaps: 7
US-10-712-629b-2 (1-746) x US-10-821-234-964 (1-1166)
QY 43 CAGTCACTGTGGGCGCGCAGGCGCAGCAGCTCAGCTCTGGGCTTGGGCTTCCCT 102
Db 105 GlnGlyProGlnGlyProLysGlyAsp-----ProGlyProGlyIlePro 120
QY 103 GGTGATCTGAGCAGCAGCAGCGCTCGGGGAAGGCGAGGTGTCCAGCTAAC----- 158
Db 121 GlyArgAsn--Gly---AspProGlyIleProGlyGlnProGlySerProGlySerProG 139
QY 159 -----AGCGGCTCAAGAACTAATGATCCA----- 182
Db 139 lyProProGlyIleCysGluSerCysProThrGlyProGlnAenTySerProGlnTyra 159
QY 183 -----GCAGTTTGGTGGC-----GGCCCAA 201
Db 159 spSerTyraAspValLysSerGlyValAlaValGlyGlyLeuAlaGlyTyProGlyProA 179
QY 202 CTGACGTGCAACAAACGCTCTTCTCCGACACGCCCAAGATCACGCCCTGGCCCTGGGC 261
Db 179 lagly----- 180
QY 262 GCAGACCCCGCTCCGAGATGCCCGCCAGCAACGCTTGGCCCACTTCACGAGCTGGCC 321
Db 181 -----ProProGlyProProGlyProGlyThr-----SerGlyH 193
QY 322 ATCATCTAGTCAGGAGATCGTGACTTCGCTTAAGCAAGTGCTGTTTCTTCGACAGCTG 381
Db 193 isProGlySerProGlySerProGlyTyTrGlnGlyProGlyGluProGlyGlnAlaG 213
QY 382 GCGCGGAGAGACAGATCCCTCTGAAGGATCCATTCAGATCATGTGCTGCTAGAG 441
Db 213 lyProSerGlyProProGlyProProGlyAlaIleGlyProSerGlyProAlaGlyLysA 233
QY 442 ACAGCAGCGCGCTACAAACA 461
Db 233 spGlyGluSerGlyArgPro 239

RESULT 8
US-11-053-100-48
; Sequence 48, Application US/11053100
; Publication No. US20050255554A1
; GENERAL INFORMATION:
; APPLICANT: CHILKOTI, Ashutosh
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION
; FILE REFERENCE: 4176-101 CIP
; CURRENT APPLICATION NUMBER: US/11/053,100
; CURRENT FILING DATE: 2005-02-08
; PRIOR APPLICATION NUMBER: US 09/812,382
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,659
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 48
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(624)
; OTHER INFORMATION: pET15b-SD3-ELP1-60-throm-Batrogen receptor ligand binding domain
US-11-053-100-48

Alignment Scores:
Pred. No.: 0.0455 Length: 624
Score: 111.00 Matches: 39

Percent Similarity:	40.97%	Conservative:	20
Best Local Similarity:	27.08%	Mismatches:	61
Query Match:	8.08%	Indels:	24
DB:	7	Gaps:	4
US-10-712-629B-2 (1-746) x US-11-053-100-49 (1-624)			
QY	52	GTGGGGCCGAGGCGAGCAGCAGCTCAGCCTCTTGGG---CCTGGGGCTTCC---CCTGGT	105
DB	284	ValGlyValProGlyValGlyValProGlyValProGlyValProGlyValProGlyValProGly	303
QY	106	GGATCTGAGGCGAGCAGGCTCCGGGAAGCGAGGT-----	147
DB	304	GlyGlyValProGlyTrpProSerGlyGlyGlySerIleGlyProLeuVal	323
QY	148	-----GTCCAGCTAACAGCGGCTCAAGAACTAATGATCCAGCAG	186
DB	324	ProArgGlySerHisMetSerLysLysAenSerLeuAlaLeuSerLeuThrAlaAspGln	343
QY	187	TTGGTGGCGGCCCAACTGCAGTGCAAAACGCTCTTCCGACAGCCCAAGTCACG	246
DB	344	MetValSerAlaLeuLeuAspAlaGluProProIleLeuTy-SerGlu	359
QY	247	CCCTGGCCCTGGGCGCAGACCCCGAGTCCCGAGATGCCCGCAGCAACGCTTGGCCAC	306
DB	360	-----TyrAspProThrArgProPheSerGluAlaSerMetMetGlyLeu	374
QY	307	TTACGAGGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGTAAGCAAGTGCCT	366
DB	375	LeuThrAsnLeuAlaAspArgGluLeuValHisMetIleAsnTrpAlaLysArgValPro	394
QY	367	GTTTCTCAGCTGGGCGGAGGAGCAGATCGCCCTCTCGAAGCATCCACTATCGAG	426
DB	395	GlyPheValAspLeuThrLeuHisAspGlnValHisLeuLeuGluCysAlaTrpLeuGlu	414
QY	427	ATCATGCTGCTA 438	
DB	415	IleLeuMetIle 418	
RESULT 9			
US-11-053-100-49			
; Sequence 49, Application US/11053100			
; Publication No. US2005025554A1			
; GENERAL INFORMATION:			
; APPLICANT: CHILKOTI, Ashutosh			
; TITLE OF INVENTION: FUSION PEPTIDES ISOLATABLE BY PHASE TRANSITION			
; FILE REFERENCE: 4176-101 CIP			
; CURRENT APPLICATION NUMBER: US/11/053,100			
; CURRENT FILING DATE: 2005-02-08			
; PRIOR APPLICATION NUMBER: US 09/812,382			
; PRIOR FILING DATE: 2001-03-20			
; PRIOR APPLICATION NUMBER: US 60/190,659			
; PRIOR FILING DATE: 2000-03-20			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: Patent in version 3.3			
; SEQ ID NO 49			
; LENGTH: 774			
; TYPE: PRT			
; ORGANISM: Artificial			
; FEATURE:			
; OTHER INFORMATION: Synthetic Construct			
; FEATURE:			
; NAME/KEY: MISC FEATURE			
; LOCATION: (1)-(774)			
; OTHER INFORMATION: pET15b-SD5-ELPI-90-throm-Estrogen receptor ligand binding domain			
US-11-053-100-49			
Alignment Scores:		Length:	774
Pred. No.:	0.0452	Matches:	39
Score:	111.00	Conservative:	20
Percent Similarity:	40.97%	Mismatches:	61
Best Local Similarity:	27.08%	Indels:	24
Query Match:	8.08%		
US-10-712-629B-2 (1-746) x US-11-059-982-1 (1-1560)			
QY	43	CAGTCACTGTGGGCGCGAGGCGCAGCAGCTCAGCCTCTGGGCT-----	90

```

Db 256 GlnValProVal-----SerGlyTyrLeuAlaSerAlaAlaGlyProSerGluProVal 273
QY 91 -----GGGCTTCCCT----- 102
Db 274 ThrLeuAlaSerAlaGlyValSerProGlnGlyAlaGlyLeuValIleGlnIysAenLeu 293
QY 103 -----GGTGATCTGAGGCGAGC 120
Db 294 SerAlaAlaValAlaThrThrLeuAenGlyAenSerValPheGlyGlyAlaGlyAlaAla 313
QY 121 AGCCAGGCTCCGGGAGGCGGCTCCAGCTTAACAGCGCTCAAGAACTAATGATC 180
Db 314 SerAlaProThrGlyThrProSerGlyGlnProLeuAlaValAlaProGlyLeuGlySer 333
QY 181 CAGCAGTTGGTG---CGCGCCCAACTGCAGTGCACAAACCGCTCTCTCCACAGCC 237
Db 334 SerProLeuValProAla-ProAenValIleuHisArgThrProThrProIleGlnPr 353
QY 238 AAAGTCAGC-----CTGGCCCTGGGGCGCAGACCCCGCTCCGAGATGCC 285
Db 353 olysProAlaGlyValLeuProProLysLeuTyrGlnLeuThrProLysProPheAlaPr 373
QY 286 CCGCAGCAACGCTTTCGCCCATCTCAGGAGCTGGCCATCATCTCAGTCCAGGAGATCGTG 345
Db 373 oAlaGlyAlaThrLeu---ThrIleGlnGlyGluProGlyAlaLeuProGlnProLy 392
QY 346 GACTTCGCTAAGCAAGTGCCTGTTCTCTGAGCTGGCGCGGAG-----AC 393
Db 392 sAlaProGlnAenLeuThrPheMetAlaAlaGlyysAlaGlyGlnAenValValLeuSe 412
QY 394 CAGATGGCCCTCCTGAAGCATCTCACTATFCGAGATCATGCTCTAGACACCCAGCGC 453
Db 412 tGlyPheProAla-----ProAlaLeuGlnAlaAenValPheLysGlnProAl 429
QY 454 TACACACGACAGACAGATGATACCTCTCTGAGACTTCACCTACAGCAAGGAGACT 513
Db 429 aThrThrThr-----GlyAlaAl 435
QY 514 TCACCGCTGACGGCTCAGGTGGAGTTCATCAACC-----CATCTTCGAGTTCT 564
Db 435 aProPro---GlnProProGlyAlaLeuSerLysProMetSerValHisLeu----- 451
QY 565 CGCGGGCCATCGCGCGCTGGCGCTGAGCAGCCTGAGTACGCCCTCTCATCGCCATCA 624
Db 452 -----LeuAenGlnGlySerSerIleValIleProAla-----Gl 463
QY 625 ACATCTTCTC-----GCCCGACCGGC----- 646
Db 463 nHisMetLeuProGlyGlnAenGlnPheLeuLeuProGlyAlaProAlaValGlnLeuPr 483
QY 647 -CAACGTCAGGAGCGCGCGCGCT-----GGAGGC-----GTTGCAGCAGC 687
Db 483 oGlnGlnLeuSerAlaLeuProAlaAenValGlyGlyGlnIleLeuAlaAlaAlaPr 503
QY 688 CTAAGTGGAGCGCTGCTGCTCTACAC-----GGCATCAAGAGCGCGCAGGACC 738
Db 503 oHisThrGlyGlyGlnLeuIleAlaAenProIleLeuThrAenGlnAenLeuAlaGlyPr 523
QY 739 A 739
Db 523 o 523

```

RESULT 11

```

US-10-821-234-1182
; Sequence 1182, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andermani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A

```

```

; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PC_SEQ_genes Version 1.0
; SEQ ID NO 1182
; LENGTH: 1874
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1182

Alignment Scores:
Pred. No.: 0.074 Length: 1874
Score: 108.00 Matches: 67
Percent Similarity: 33.92% Conservative: 10
Best Local Similarity: 29.52% Mismatches: 82
Query Match: 7.86% Indels: 70
DB: 1 Gaps: 12
US-10-712-629B-2 (1-746) x US-10-821-234-1182 (1-1874)

```

```

QY 49 CCTGTGGGGCCCGAGGGCAGCAGCTCA-----GCCTCTGGGCTTGGGGCTTCCCT 102
Db 1289 ProProGlyProArgGlyProSerGlyAlaProGlyAlaAspGlyProGlnGlyProPro 1308
QY 103 GGTGGA-----TCTGAGCAGCAG- 122
Db 1309 GlyGlyIleGlyAenProGlyAlaValGlyGluLysGlyGluProGlyGluAlaGlyGlu 1328
QY 123 CCAGGGCTCCGGGGAAGCGAGGTGTCCAGCTTAACAGCGGTCAAGAACTAATGATCCA 182
Db 1329 ProGlyProSerGlyArgSerGlyProProGlyProLysGlyGluArgGlyGluLysGly 1348
QY 183 GCAGTTGGTGGCGGCCAACTCAGTGCACAAACGCTCTCTCTCGACACAGCCCAAGT 242
Db 1349 Glu-----SerGlyProSerGly----- 1354
QY 243 CACGCCCTGGCCCTGGGGCGCAGACCCAGTCCCGAGATGCCCGCAGCAACGCTTTCG 302
Db 1355 -----AlaAlaGly-----ProProGlyProLysGlyProProGlyAenAspGly 1369
QY 303 CCACCTTACGAGCTGGCCATCATCTCAGTCCAGGAGATCGTGGACTTCGCTAAGCAAGT 362
Db 1370 ProLysGlySerProGlyProValGlyPheProGlyAenProGly----- 1384
QY 363 GCCTGGTTCTTCGAGCT-----GGCGGGGAGGACCAAGATCGCCCTCTGAAGCATC 416
Db 1385 -----ProProGlyGluProGlyProAlaGlyGlnAenProPro----- 1398
QY 417 CACTATCAGATCATGCTGCTAG-AGACAGCCAGCGGCTACAAACACGAGACAGAGTGA 475
Db 1399 -----GlyAenPheGlyAenAspGlyGluProGly-----GlnThrGly 1411
QY 476 TCACCTTCTTGAGGACTTCACCTACAGCAAGGACGACTTCCACCGTGCAGGCTGCAGGT 535
Db 1412 SerProGly-----ProThrGlyGluProGlyProSerGlyProProGly 1426
QY 536 GGAGTTTCATCAACCCCATCTTCGAGTTCTTCGGGGCCATTCGGCGGCTGGGCT---GGA 592
Db 1427 LysArgGlyProPro-----GlyProAlaGlyProGluGly 1438
QY 593 CGACGCTGAGTACGCCCTGCTCATCGCCATCAACATCTTCTCGCGGACCGGCCCAACGT 652
Db 1439 ArgGlnGlyGluLysGlyAlaLysGlyGluAlaGlyLeuGluGlyProProGlyLysThr 1458
QY 653 GCAGGAGCGCGCGCGCTGGA 673
Db 1459 GlyProIleGlyProGlnGly 1465

```

RESULT 12

```

US-11-053-100-50
; Sequence 50, Application US/11053100

```

1.

Search completed: November 25, 2005, 22:32:04
Job time : 19.5 secs

The Fog with (copy)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:03:31 ; Search time 165 Seconds
(without alignments)
1754.883 Million cell updates/sec

Title: US-10-712-629B-18
Perfect score: 3704
Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQEAQ 693

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main: *
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pbp: *
2: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pbp: *
3: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pbp: *
4: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pbp: *
5: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pbp: *
6: /cgn2_6/prodata1/pubpaa/US11_PUBCOMB.pbp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3682	99.4	693	US-10-452-858C-93	Sequence 93, Appl
2	3682	99.4	1182	US-10-024-368-6	Sequence 6, Appl
3	3510	94.8	1207	US-10-024-368-5	Sequence 5, Appl
4	3095.5	83.6	1189	US-10-024-368-4	Sequence 4, Appl
5	3089.5	83.4	1189	US-10-122-013-17	Sequence 17, Appl
6	3088.5	83.4	984	US-10-024-368-2	Sequence 2, Appl
7	3077.5	83.1	1189	US-10-024-368-3	Sequence 3, Appl
8	729	19.7	1265	US-10-408-765A-795	Sequence 795, App
9	728.5	19.7	2055	US-10-193-874-20	Sequence 20, Appl
10	728.5	19.7	2321	US-10-193-874-17	Sequence 17, Appl
11	728.5	19.7	2358	US-10-193-874-14	Sequence 14, Appl
12	728.5	19.7	2540	US-10-193-874-16	Sequence 16, Appl
13	728.5	19.7	2552	US-10-193-874-15	Sequence 15, Appl
14	719	19.4	2210	US-10-719-993-589	Sequence 589, App
15	714	19.3	2353	US-10-719-993-588	Sequence 588, App
16	714	19.3	2535	US-10-719-993-587	Sequence 587, App
17	700.5	18.9	1417	US-10-408-765A-1992	Sequence 1992, Ap
18	606	16.4	671	US-10-264-049-2730	Sequence 2730, Ap
19	603.5	16.3	1212	US-10-170-385-19	Sequence 19, Appl
20	550.5	14.9	854	US-11-097-143-17979	Sequence 17979, A
21	539	14.6	953	US-10-193-874-18	Sequence 18, Appl
22	448.5	12.1	1192	US-10-755-889-640	Sequence 640, App
23	313.5	8.5	872	US-10-437-963-119587	Sequence 119587,
24	297	8.0	747	US-10-425-114-38701	Sequence 38701, A
25	297	8.0	941	US-10-424-593-210074	Sequence 210074,
26	280.5	7.6	874	US-10-437-963-158893	Sequence 158893,
27	279.5	7.5	363	US-10-425-114-42461	Sequence 42461, A

ALIGNMENTS

RESULT 1

US-10-452-858C-93
; Sequence 93, Application US/10452858C
; Publication No. US20040086945A1
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadanam
; APPLICANT: Gerwe, Gina S.
; TITLE OF INVENTION: HAIRLESS PROTEIN-INTERACTING PARTNER COMPLEXES AND METHODS THERE
; FILE REFERENCE: 8956P
; CURRENT APPLICATION NUMBER: US/10/452,858C
; CURRENT FILING DATE: 2003-06-02
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-452-858C-93

Query Match		99.4%;	Score 3682;	DB 4;	Length 693;
Best Local Similarity		99.4%;	Pred. No. 1.5e-278;		
Matches 689;		Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	VTQCSCVQAAGEVGLTGH	SOKSRRSPLEEKQLEEDSSATSEGGGPGPEASLNKGL	60	
Db	1	VTQCSCVQAAGEVGLTGH	SOKSRRSPLEEKQLEEDSSATSEGGGPGPEASLNKGL	60	
QY	61	AKHLLSGLDRLCLRLKRE	ALAWAQREGQGPAMTSDSPGIPHCSCRCHHGLFNTHWC	120	
Db	61	AKHLLSGLDRLCLRLKRE	ALAWAQREGQGPAMTSDSPGIPHCSCRCHHGLFNTHWC	120	
QY	121	SHCSHRLCVACGR	IAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAEIS	180	
Db	121	SHCSHRLCVACGR	IAGAKNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAEIS	180	
QY	181	TYMHQAWAKFDIRGHCF	QVQVDARVWAPGDGQOKEPTKPTPTPQSCNGDSNRTKDIXE	240	
Db	181	TYMHQAWAKFDIRGHCF	QVQVDARVWAPGDGQOKEPTKPTPTPQSCNGDSNRTKDIXE	240	
QY	241	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKCLGHDRTHMAFAPVTPALPSDDRIT	300	
Db	241	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKCLGHDRTHMAFAPVTPALPSDDRIT	300	
QY	301	NILDSIIIAQVVERKIQE	KALGSLRGLRAGSLRGLSLPSVTRLSPPGALLWQBP RPK	360	
Db	301	NILDSIIIAQVVERKIQE	KALGSLRGLRAGSLRGLSLPSVTRLSPPGALLWQBP RPK	360	
QY	361	HGFHLFQBHWRCQPV	LVSGIQKTLRLSLMGWEALGTGGQVQSJLTALGPQPTNLDSTA	420	
Db	361	HGFHLFQBHWRCQPV	LVSGIQKTLRLSLMGWEALGTGGQVQSJLTALGPQPTNLDSTA	420	

Db 361 HGPHLFOEHWROQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 420
QY 421 FWEGFHPETRPKLDGSGVLLHRLTGLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 480
Db 421 FWEGFHPETRPKLDGSGVLLHRLTGLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 480
QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
QY 541 GLDGEGLWSPGSGTSTVHVHFRADAQRIIRRFQMVCPAGAGTLEPGAGCSYLDAGLRR 600
Db 541 GLDGEGLWSPGSGTSTVHVHFRADAQRIIRRFQMVCPAGAGTLEPGAGCSYLDAGLRR 600
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
Db 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
QY 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693
Db 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693

RESULT 2

US-10-024-368-6
; Sequence 6, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mouse
US-10-024-368-6

Query Match 99.4%; Score 3682; DB 4; Length 1182;
Best Local Similarity 99.4%; Pred. No. 3,2e-278;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKOLEEDSSATSEEGGGGPGPEASLNKGL 60
Db 490 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKOLEEDSSATSEEGGGGPGPEASLNKGL 549
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGPHCCSRCHHGLFNTHWRC 120
Db 550 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGPHCCSRCHHGLFNTHWRC 609
QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQAGHAACSLILTOFVSSQALAEIS 180
Db 610 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQAGHAACSLILTOFVSSQALAEIS 669
QY 181 TVNHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 240
Db 670 TVNHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 729
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 730 ETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789
QY 301 NILDSIIAQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPPGALLMLQEPKPK 360
Db 790 NILDSIIAQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPPGALLMLQEPKPK 849

QY 361 HGPHLFOEHWROQOPVLVSGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
Db 850 HGPHLFOEHWROQOPVLVSGIQKTLRLSLWGMALGTLGGQVQTLTALGPPQPTNLDSTA 909
QY 421 FWEGFHPETRPKLDGSGVLLHRLTGLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 480
Db 910 FWEGFHPETRPKLDGSGVLLHRLTGLGDKDASRVQNLVSSLPPEYCAHQGKLNLSYLP 969
QY 481 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 970 LGTLHPLPQLWAAAYGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 1029
QY 541 GLDGEGLWSPGSGTSTVHVHFRADAQRIIRRFQMVCPAGAGTLEPGAGCSYLDAGLRR 600
Db 1030 GLDGEGLWSPGSGTSTVHVHFRADAQRIIRRFQMVCPAGAGTLEPGAGCSYLDAGLRR 1089
QY 601 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 660
Db 1090 RLREEWGVSCTWLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHOG 1149
QY 661 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 693
Db 1150 ASLPPDHRMLYAQMRAVFOAVKAAVGALEAK 1182

RESULT 3

US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-10-024-368-5

Query Match 94.8%; Score 3510; DB 4; Length 1207;
Best Local Similarity 95.1%; Pred. No. 9e-265;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKOLEEDSSATSEEGGGGPGPEASLNKGL 60
Db 517 VTQCSCVQAAGEVGLTGHSSQKRRSPLEEKOLEEDSSATSEEGGGGPGPEASLNKGL 574
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGPHCCSRCHHGLFNTHWRC 120
Db 575 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPANTEDSPGPHCCSRCHHGLFNTHWRC 634
QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQAGHAACSLILTOFVSSQALAEIS 180
Db 635 SHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQAGHAACSLILTOFVSSQALAEIS 694
QY 181 TVNHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 240
Db 695 TVNHQAWAKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPPQPSGNGDSNRTKDIKE 754
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 755 ETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 814
QY 301 NILDSIIAQVVERKIQEALGPGIRAGSLRKGSLPLSPVTRTSLSPPGALLMLQEPKPK 360

Db 815 NILDSTIIAQVVERKIQEAKLPGRLRAGSLRKLGLSLPLSPVTRTSLPPGALLWLQEPK 874
QY 361 HGPHLFOEHWROQPVLVSGIQKTLRLSLWGMALGTLGQVQSLTALGPQPTNLDSTA 420
Db 875 HGPHLFOEHWROQPVLVSGIQKTLRLSLWGMALGTLGQVQSLTALGPQPTNLDSTA 934
QY 421 FMEGFSHPETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 935 FMEGFSHPETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 994
QY 481 LGTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAELPPEYCAHOGKLNLSYLP 540
Db 995 LGTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAELPPEYCAHOGKLNLSYLP 1054
QY 541 GLDGEGLWSPGTSQTSVWHVFRQAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 600
Db 1055 GLDGEGLWSPGTSQTSVWHVFRQAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAGLR 1114
QY 601 LRREWGVSCWTLLOAPGSAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACLCHQ 660
Db 1115 LRREWGVSCWTLLOAPGSAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACLCHQ 1174
QY 661 ASLPPDHRMLYAQMDDRAVFAVKAAGALQEA 693
Db 1175 ASLPPDHRMLYAQMDDRAVFAVKAAGALQEA 1207

RESULT 4
US-10-024-368-4
; Sequence 4, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/024,368
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-04-07
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-4

Query Match 83.6%; Score 3095.5; DB 4; Length 1189;
Best Local Similarity 83.2%; Pred. No. 2.1e-232;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCQSCVQAAGEVGLTGHKSRRSPL-EEKQLEEDSSATSEEGGGGPGPEASLNKG 59
Db 494 LAQCQSCAQAAGEGGCHACHSQVRRSPGLGELQEEEDTATNSSSEEGGSGPDSRLSTG 553
QY 60 LAKHLLSGLDRLCLRLKEREALAWAQEGOGPANTEPSGIPHCSCCHHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLRLKEREALAWAQEGOGPANTEPSGIPHCSCCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRITAGAGNREKTSQEQHTDDCAQEAAGHAACSLILTFVSSQALAE 179
Db 614 CPCSRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLILTFVSSQALAE 673
QY 180 STMVHQAQKFDIRGHCFQVDARVWAPGDGGQOKEPTKPTTPOPCNGSDNRKTDIK 239
Db 674 STAMHQVWVVFDIRGHCFQVDARVWAPGDGGQOKEPTKPTTPOPCNGSDNRKTDIK 733
QY 240 BETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Db 734 BETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793

QY 300 TNILDSIIAQVVERKIQEAKLPGRLRAGSLRKLGLSLPLSPVTRTSLPPGALLWLQEPK 358
Db 794 TNILDSIIAQVVERKIQEAKLPGRLRAGSLRKLGLSLPLSPVTRTSLPPGALLWLQEPK 853
QY 359 -PKHGHPLFOEHWROQPVLVSGIQKTLRLSLWGMALGTLGQVQSLTALGPQPTNLD 417
Db 854 CPMRGPHLFOEHWROQPVLVSGIQKTLRLSLWGMALGTLGQVQSLTALGPQPTNLD 913
QY 418 STAFWEGFSPETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLS 477
Db 914 STTFWEGFSPETPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLS 973
QY 478 YLPLGLTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAELPPEYCAHOGKLN 537
Db 974 YLPLGLTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAELPPEYCAHOGKLN 1033
QY 538 FLISGLDGEGLWSPGTSQTSVWHVFRQAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAG 597
Db 1034 FLISGLDGEGLWSPGTSQTSVWHVFRQAQRIIRRFQMVCPAGAGTLEPGAGSCYLDAG 1093
QY 598 LRRLREWGVSCWTLLOAPGSAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACL 657
Db 1094 LRRLREWGVSCWTLLOAPGSAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSACL 1153
QY 658 HQGASLPPDHRMLYAQMDDRAVFAVKAAGALQEA 693
Db 1154 HQGASLPPDHRMLYAQMDDRAVFAVKAAGALQEA 1189

RESULT 5
US-10-122-013-17
; Sequence 17, Application US/10122013
; Publication No. US20030077614A1
; GENERAL INFORMATION:
; APPLICANT: Christiano, Angela
; TITLE OF INVENTION: NUCLEIC ACIDS FOR INHIBITING HAIRLESS PROTEIN EXPRESSION AND MET
; FILE REFERENCE: US/10/122,013
; CURRENT FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-122-013-17

Query Match 83.4%; Score 3089.5; DB 4; Length 1189;
Best Local Similarity 83.0%; Pred. No. 6.3e-232;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCQSCVQAAGEVGLTGHKSRRSPL-EEKQLEEDSSATSEEGGGGPGPEASLNKG 59
Db 494 LAQCQSCAQAAGEGGCHACHSQVRRSPGLGELQEEEDTATNSSSEEGGSGPDSRLSTG 553
QY 60 LAKHLLSGLDRLCLRLKEREALAWAQEGOGPANTEPSGIPHCSCCHHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLRLKEREALAWAQEGOGPANTEPSGIPHCSCCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRITAGAGNREKTSQEQHTDDCAQEAAGHAACSLILTFVSSQALAE 179
Db 614 CPCSRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGHAACSLILTFVSSQALAE 673
QY 180 STMVHQAQKFDIRGHCFQVDARVWAPGDGGQOKEPTKPTTPOPCNGSDNRKTDIK 239
Db 674 STAMHQVWVVFDIRGHCFQVDARVWAPGDGGQOKEPTKPTTPOPCNGSDNRKTDIK 733
QY 240 BETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 299
Db 734 BETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPSDDRI 793
QY 300 TNILDSIIAQVVERKIQEAKLPGRLRAGSLRKLGLSLPLSPVTRTSLPPGALLWLQEPK 358

Db 794 TNLDSIIAQVVERKIQEALGFLRAGPLRGLGLPLSPVRRPLPPPGALLWLQEPQP 853
QY 359 -PKHGPHLFOEHWROQOPVLVSGIQKTLRLSLMGMEALGTGGOVQSLTALGPPQPTNLD 417
Db 854 CPRRGPHLFOEHWROQOPVLVSGIQRTLOGNLWTEALGALGGQVQALSPLGPPQPSL 913
QY 418 STAFWEGFSPHETPRPKLDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 477
Db 914 STTFWEGFSPHETPRPKLDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 973
QY 478 YLPLGLTLPLPOLWAAVGNVSHRGHLGTLKNCVSDLSILVHAQAQLPWPYRAQKD 537
Db 974 YLPPGLALRPLPOLWAAVGNVSHRGHLGTLKNCVSDLSILVHAQAQLPWPYRAQKD 1033
QY 538 FLISGLDEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSGCYLDAG 597
Db 1034 FLISGLDEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSGCYLDAG 1093
QY 598 LRRRLREEMVSGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 657
Db 1094 LRRRLREEMVSGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVCAALQEA 693
Db 1154 HQGASLPDPHRLMYAQMDRAVFOAVKAAVCAALQEA 1189

RESULT 6

US-10-024-368-2
; Sequence 2, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-2

Query Match 83.4%; Score 3088.5; DB 4; Length 984;
Best Local Similarity 83.0%; Pred. No. 5.8e-232;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;

QY 1 VTQCSCVQAGVGVLTGHSQKRRSPL-EEKLEEDSDSATSEEGGGGPGPEASLNGK 59
Db 289 LAQCSCAAAGGEGGHACHSQVRRSPLOGELQOEEDTATNSSSEEGGPGPSRLSTG 348
QY 60 LAKHLISGLGDLRLLREREAALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 119
Db 349 LAKHLISGLGDLRLLREREAALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRIAGAGKNNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 179
Db 409 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTOFVSSQALAE 468
QY 180 STVHQAWKAFDTRGHCFQVDARVWAPGQGGQKEPTKPTTPOPSNGDSNRKDIK 239
Db 469 STAMHQVWKFDRGHCFQVDARVWAPGQGGQKEPTKPTTPOPSNGDSNRKDIK 528
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 299
Db 529 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 588

QY 300 TNLDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSVRTRLSPPGALLWLQEPQP- 358
Db 589 TNLDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSVRTRLSPPGALLWLQEPQP 648
QY 359 -PKHGPHLFOEHWROQOPVLVSGIQKTLRLSLMGMEALGTGGOVQSLTALGPPQPTNLD 417
Db 649 CPRRGPHLFOEHWROQOPVLVSGIQRTLOGNLWTEALGALGGQVQALSPLGPPQPSL 708
QY 418 STAFWEGFSPHETPRPKLDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 477
Db 709 STTFWEGFSPHETPRPKLDEGSVLLHRLTGDKDASRVQNLVSSLPLPEYCAHOGKLNLAS 768
QY 478 YLPLGLTLPLPOLWAAVGNVSHRGHLGTLKNCVSDLSILVHAQAQLPWPYRAQKD 537
Db 769 YLPPGLALRPLPOLWAAVGNVSHRGHLGTLKNCVSDLSILVHAQAQLPWPYRAQKD 828
QY 538 FLISGLDEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSGCYLDAG 597
Db 829 FLISGLDEGLWSPGSGTSTVWHVFRADQRIIRRFQWVCPAGAGTLEPGAGSGCYLDAG 888
QY 598 LRRRLREEMVSGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 657
Db 889 LRRRLREEMVSGVSWTLLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQALC 948
QY 658 HQGASLPDPHRLMYAQMDRAVFOAVKAAVCAALQEA 693
Db 949 HQGASLPDPHRLMYAQMDRAVFOAVKAAVCAALQEA 984

RESULT 7

US-10-024-368-3
; Sequence 3, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/10/024,368
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-368-3

Query Match 83.1%; Score 3077.5; DB 4; Length 1189;
Best Local Similarity 82.9%; Pred. No. 5.4e-231;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAGVGVLTGHSQKRRSPL-EEKLEEDSDSATSEEGGGGPGPEASLNGK 59
Db 494 LAQCSCAAAGGEGGHACHSQVRRSPLOGELQOEEDTATNSSSEEGGPGPSRLSTG 553
QY 60 LAKHLISGLGDLRLLREREAALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 119
Db 554 LAKHLISGLGDLRLLREREAALAWAQREGQGPAMTDSFGIPHCSCRCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRIAGAGKNNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLMLTOFVSSQALAE 673
QY 180 STVHQAWKAFDTRGHCFQVDARVWAPGQGGQKEPTKPTTPOPSNGDSNRKDIK 239
Db 674 STAMHQVWKFDRGHCFQVDARVWAPGQGGQKEPTKPTTPOPSNGDSNRKDIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDIRIHMAFAPVTPALPSDDR 299

Db 734 EETPDSAEETAEADRAGRPFCPSCELLASTAVKLCGHDRHMAPVTPALPSDDRI 793
Qy 300 TNLDSIIAOWVERKIOEKALGCLAGSLRGLSLPLSVTRLSPPCALLWLOEPR- 358
Db 794 TNLDSIIAOWVERKIOEKALGCLAGSLRGLSLPLSVTRLSPPCALLWLOEPR 853
Qy 359 -PKHGFLFOEHWRRQGPVLVSGIQKTLRLSLMGMEALGTGGQVQSLSLTALGPPQPTNLD 417
Db 854 CPRRGFLFOEHWRRQGPVLVSGIQKTLRLSLMGMEALGTGGQVQSLSLTALGPPQPTNLD 913
Qy 418 STAFWEGFSPHETRPKLDDESVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 477
Db 914 STTFWEGFSPHETRPKLDDESVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 973
Qy 478 YLPLGLTLHLEPQLMAAYGVNSHRGLGKNCVSDLSILVHAEALPWPYRAQKD 537
Db 974 YLPPGLALRLEPQLMAAYGVNSHRGLGKNCVSDLSILVHAEALPWPYRAQKD 1033
Qy 538 FLSGLDGEGLSVSGSQTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAG 597
Db 1034 FLSGLDGEGLSVSGSQTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAG 1093
Qy 598 LRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACL 657
Db 1094 LRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACL 1153
Qy 658 HQGASLPPDHRMLYAQMDRAVFAVKAACALQ 693
Db 1154 HQGASLPPDHRMLYAQMDRAVFAVKAACALQ 1189

RESULT 8
US-10-408-765A-795
; Sequence 795, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 795
; LENGTH: 1265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-795

Query Match 19.7%; Score 729; DB 4; Length 1265;
Best Local Similarity 29.2%; Pred. No. 1.9e-47;
Matches 219; Conservative 117; Mismatches 270; Indels 144; Gaps 25;
Qy 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTDESPGPHCCSRCHHGLFNTHWRC 120
Db 528 SKYILDIIIGDKFCQLVTSEKTSLSWVKDAK-IWKRAVRGVREMCDAEATLNFNIHWVC 586
Qy 121 SHCSRHLVACGRIAGAGKNRKTSQEQHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
Db 587 QKGFVVCCLDCYK----AKERKSSRDKELYAMKVKQGPQHDHKLHMLPTQIIIPGSLVTL 642
Qy 180 STVMHQAQAFDIRGHCF-----QVDARVWAPGDGQQ-----KEPTE 218
Db 643 LDAMHTLRKYGIKSHCHTCKNQNLQVGNFPTMNGVSQVLQNLVNLHNSKISCLMPESQQ 702

Qy 219 KTFPTQPSCNGSDNRKTIKEE---TPDSTESP----- 249
Db 703 NTFPKSEK-NGSSPESDVGTDNKLTPESQSPHLWLADLAQKAREEKKENKELTLEN 760
Qy 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCGHDRHMAPV- 289
Db 761 QIKEREQDNSESPNGRTSPLVSQNNQEGSTLRLDTTACKLRVSGTDAIAPAPVYSM 820
Qy 290 -TPALSDDDRITNLDIIIAOVVERKI-----QEKALGPGURAGSLRKGL 334
Db 821 GAPSSKSGTMTNLDIIIASVVVENKIPPSKTSKINVKPELKEEPEESIISAVENNKIYS 880
Qy 335 SLPLSVTRLSPPGALLMLQEPKPHGFLFOEHWRRQGPVLVSGIQKTLRLSLMGMEA 394
Db 881 DIPHSWICEK-----HILMLKDYKNSSNMKLKPECKKQGPAPVSVGHKKNISLWKAES 935
Qy 395 LG-TLGGQVQSLSLTALGPPQPTNLDSTAFWEGFSPHETRPKLDGSGVLLHRTLGDKD- 450
Db 936 ISLDFGDHQAADLLNCKDSIISANVKEFWDFGFEVSKRQKNKSGETVVLKL-----KOWPS 991
Qy 451 -----ASRVQNLVSSLPPEYCAHOGKUNLASVLPGLTLHP-LPEPOLWAAYGVNSH 501
Db 992 GEDFTKMPARVEDLLKSLPLPEYCNPEGKFNASHLP-GFFVRPDLGPRCLCSAYGVAA 1050
Qy 502 RGH-LGTKMLCNEVSDLSILVHAEALPWPYRAQKDFLSGLDGEGLWS----- 549
Db 1051 KOHDIGTTLNLTIEVDVNLVYVGIAGKNGILSKAGILKKPEEDLDDILRKLDKDSSE 1110
Qy 550 -PGSQTSTVHVFRAQDAQRIIRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRLRE 604
Db 1111 IPGA-----LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPIRDQSWYVYVKKLRQLLE 1166
Qy 605 EWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSACLCHQASLP 664
Db 1167 EYGVRTCTLIQFLGDAIVPALGALHQVQNFHSCIQVTEDFVSPEHLVSPHFLTQELRLK 1226
Qy 665 P-----DHRMLYAQMDRAVFAVKAACALQ 690
Db 1227 BEINYDDKL---QVKNILYHAKVMYRALX 1253

RESULT 9
US-10-193-874-20
; Sequence 20, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-20

Query Match 19.7%; Score 728.5; DB 4; Length 2055;
Best Local Similarity 29.1%; Pred. No. 4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
Qy 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTDESPGPHCCSRCHHGLFNTHWRC 120
Db 1317 SKYILDIIIGDKFCQLVTSEKTSLSWVKDAK-IWKRAVRGVREMCDAEATLNFNIHWVC 1375
Qy 121 SHCSRHLVACGRIAGAGKNRKTSQEQHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
Db 1376 QKGFVVCCLDCYK----AKERKSSRDKELYAMKVKQGPQHDHKLHMLPTQIIIPGSLVTL 1431
Qy 180 STVMHQAQAFDIRGHCF-----QVDARVWAPGDGQQ-----KEPTE 218

Db 1432 LDAMHTLRREYGIKSHCHCTNKQNLQVGNPFTMNGVSQVLQNVLNHNSKISLCMPESQOQ 1491
 Qy 219 KTPPTPQPCNGSDNRKDIKEE---TPDSTESP----- 249
 Db 1492 NTPPKSEK---NGSSPESDVGTDNKLTPPESQSPHLWADLAQKAREEKENKELTLEN 1549
 Qy 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCIGHDRIHMAFAPV--- 289
 Db 1550 QIKEEREQDNSESPNGRTSPLVSONNEQGSTLRDLTLTTAGKLRVGVSTDAGIAFAPVYSM 1609
 Qy 290 -TPALPSDDRIITNLDIISIAOVVERKI-----QBKALGPGRLRAGSGLR 331
 Db 1610 GAPSSKSGRTMNPILDDIIASVVENKIPPSKTSKINVKPELKEEPEESIIASVDENNKLY 1669
 Qy 332 KGLSLPLSPVRTRLSPGALLWLQEPKPHGPHLFOEHRMQOPVLVSGIQKTLRLSLMG 391
 Db 1670 S--DIPHSWICEK-----HILWLKDYKNSSNNWKLFECKWQGPAPVVGHVHKKMNIWLK 1722
 Qy 392 MEALG-TLGGQVQSLTALGPQPTNLDSTAFWEGFHPETRPKLDGEGSVLLHRTLGDKD 450
 Db 1723 AESISLDFGDHQAADLLNCKDSIISANVKEFDGFEVSKRQKSGETVVLKL-----KD 1778
 Qy 451 -----ASRVQNLVSSLPPEYCAHQKLNLASYLPLGLTLHP-LLEPOLWAAYGV 498
 Db 1779 WPSGDFKTMMPARYEDLLKSLELPEYCNPEGKFNLAHLUP-GFFVRPDLGRLCSAYGV 1837
 Qy 499 NSHRGH-LGTKNLCVEVSDIISILVHAEALQPPWYRAQKDFLSGLDGEGLMS----- 549
 Db 1838 VAAKDHIDGTTNLHIEVSDVNVILVYVGIAGNGLSKAGILKKFEEDLDDILRKLKD 1897
 Qy 550 -----PGSQTSTVHVHFRQAQRIIRPLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 1898 SSEIPGA-----LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPIRDQOSWYVYNNKLQR 1953
 Qy 602 LREEMGVSWTLLQAPGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSQAOLCHOQA 661
 Db 1954 LLEEYGVTRCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESFHLTQELR 2013
 Qy 662 SLPP-----DHRMLYAQMRAVFOAKAAGALQ 690
 Db 2014 LKKEEINYDKL---QVKNILYHAVKEMVRALK 2043

RESULT 10
 US-10-193-874-17
 ; Sequence 17, Application US/10193874
 ; Publication No. US20030064396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jenapharm GmbH & Co. KG
 ; TITLE OF INVENTION: Comodulators of Nuclear Receptors
 ; FILE REFERENCE: 52145
 ; CURRENT APPLICATION NUMBER: US/10/193,874
 ; CURRENT FILING DATE: 1998-11-18
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 17
 ; LENGTH: 2321
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-193-874-17

Query Match 19.7%; Score 728.5; DB 4; Length 2321;
 Best Local Similarity 29.1%; Pred. No. 4.7e-47;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;

Qy 61 AKHLLSGDLRCLRLKRLKEREALAWAQREGGQPMATEDSPGIPHCCSRCHHGLFNTHWRC 120
 Db 1583 SKYILDIDGKFCQLVTSKLTALSWMKKDAK-IAWKRAVRGVREMCACEATLNFNIHWVC 1641
 Qy 121 SHCSHRLVCACGRIAGAKNREKTSQSOHT-DDCAQEAAGHAACSLIITQFVSSQALAE 179
 Db 1642 QKCGFVCLDCK-----AKERKSRDKELYAMWCKVQKQPHDKHLMPTQIIPGSVLTDL 1697

Qy 180 STVMQAWAKFDIRGHCF-----QVDARVWAPGDGQO-----KEPTS 218
 Db 1698 LDAMHTLRREYGIKSHCHCTNKQNLQVGNPFTMNGVSQVLQNVLNHNSKISLCMPESQOQ 1757
 Qy 219 KTPPTPQPCNGSDNRKDIKEE---TPDSTESP----- 249
 Db 1758 NTPPKSEK---NGSSPESDVGTDNKLTPPESQSPHLWADLAQKAREEKENKELTLEN 1815
 Qy 250 -----AEDGAGR-SPLPC-----PSCELLASTAVKLCIGHDRIHMAFAPV--- 289
 Db 1816 QIKEEREQDNSESPNGRTSPLVSONNEQGSTLRDLTLTTAGKLRVGVSTDAGIAFAPVYSM 1875
 Qy 290 -TPALPSDDRIITNLDIISIAOVVERKI-----QBKALGPGRLRAGSGLR 331
 Db 1876 GAPSSKSGRTMNPILDDIIASVVENKIPPSKTSKINVKPELKEEPEESIIASVDENNKLY 1935
 Qy 332 KGLSLPLSPVRTRLSPGALLWLQEPKPHGPHLFOEHRMQOPVLVSGIQKTLRLSLMG 391
 Db 1936 S--DIPHSWICEK-----HILWLKDYKNSSNNWKLFECKWQGPAPVVGHVHKKMNIWLK 1988
 Qy 392 MEALG-TLGGQVQSLTALGPQPTNLDSTAFWEGFHPETRPKLDGEGSVLLHRTLGDKD 450
 Db 1989 AESISLDFGDHQAADLLNCKDSIISANVKEFDGFEVSKRQKSGETVVLKL-----KD 2044
 Qy 451 -----ASRVQNLVSSLPPEYCAHQKLNLASYLPLGLTLHP-LLEPOLWAAYGV 498
 Db 2045 WPSGDFKTMMPARYEDLLKSLELPEYCNPEGKFNLAHLUP-GFFVRPDLGRLCSAYGV 2103
 Qy 499 NSHRGH-LGTKNLCVEVSDIISILVHAEALQPPWYRAQKDFLSGLDGEGLMS----- 549
 Db 2104 VAAKDHIDGTTNLHIEVSDVNVILVYVGIAGNGLSKAGILKKFEEDLDDILRKLKD 2163
 Qy 550 -----PGSQTSTVHVHFRQAQRIIRPLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 2164 SSEIPGA-----LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPIRDQOSWYVYNNKLQR 2219
 Qy 602 LREEMGVSWTLLQAPGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSQAOLCHOQA 661
 Db 2220 LLEEYGVTRCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESFHLTQELR 2279
 Qy 662 SLPP-----DHRMLYAQMRAVFOAKAAGALQ 690
 Db 2280 LKKEEINYDKL---QVKNILYHAVKEMVRALK 2309

RESULT 11
 US-10-193-874-14
 ; Sequence 14, Application US/10193874
 ; Publication No. US20030064396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jenapharm GmbH & Co. KG
 ; TITLE OF INVENTION: Comodulators of Nuclear Receptors
 ; FILE REFERENCE: 52145
 ; CURRENT APPLICATION NUMBER: US/10/193,874
 ; CURRENT FILING DATE: 1998-11-18
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 2358
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-193-874-14

Query Match 19.7%; Score 728.5; DB 4; Length 2358;
 Best Local Similarity 29.1%; Pred. No. 4.8e-47;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;

Qy 61 AKHLLSGDLRCLRLKRLKEREALAWAQREGGQPMATEDSPGIPHCCSRCHHGLFNTHWRC 120
 Db 1620 SKYILDIDGKFCQLVTSKLTALSWMKKDAK-IAWKRAVRGVREMCACEATLNFNIHWVC 1678
 Qy 121 SHCSHRLVCACGRIAGAKNREKTSQSOHT-DDCAQEAAGHAACSLIITQFVSSQALAE 179

Db 1679 QKQFVVCLDYK-----AKERKSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1734
QY 180 STVMQAWAKFDIRGHCF-----QVDARVWAPDGGQ-----KEPTE 218
Db 1735 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTMNGVSQVLQNLVNLHNSKISLCMPESQOQ 1794
QY 219 KTPPTPOPSCNGSDNRTKIKER---TPDSTESP----- 249
Db 1795 NTPPKSEK--NGSSPESDVTGDKLTPPESQSPHLADLABQKAREEKENKELTLEN 1852
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHRHMAFAPV--- 289
Db 1853 QIKEREQDNSESPNGRTSPLVSQNNQEGSTLRDLTTTAGKLRVGSTDAGIAFAPVYSM 1912
QY 290 -TPALPSDDRIITNLDIIIAQVVERKI-----QBKALGPGLRAGSGLR 331
Db 1913 GAPSSKSGRTMPNILDIIASVVENKI PPSKTSKINVKPELKEEPESIIISAVDENNKLY 1972
QY 332 KGLSLPLSPVTRTLSPGALLMLQEPKPKGHFLFOEHRWQOPVLVSGIQKTLRLSLWG 391
Db 1973 S--DIPHSWICEK-----HILMKDYKNSNNWKLKFCWKQGOPAVVSGVHKOMNISLWK 2025
QY 392 MEALG--TLGGQVQSLTALGPPQPTNLDSTAFWEGFSHPETRPKLDGSLVLLHRTLGDKD 450
Db 2026 AESSISLDFGDHQAADLLNCKDSIISANVKEFMDGFEEVSKRQKNSGETVWLK-----KD 2081
QY 451 -----ASRVNQLVSSLPPEYCAHQKLNLASVPLGLTLHP--LEPOLMAAYGV 498
Db 2082 WPSGDFKTMPPARYEDLLKSLPLPEYCNPEGKFNLAHLP--GFFVRPDLGPRLC SAYGV 2140
QY 499 NSHRGH--LGTKNLCVSDLSILVHAEAQLPPWYRAQKDFLSGLDGEGLWS----- 549
Db 2141 VAAKHODIGTTLNHLIEVSDVNNILVYVGIAGKNGILSKAGILKPFEEEDDDILRKRLKD 2200
QY 550 -----PGSQSTVWHVFRADAQRIIRFLQWVC--PAGAGTL---EPGAPGSCYLDAGLRRR 601
Db 2201 SSEIFGA---LWHIYAGKDVKIREFLQKISKEQGLEVLPEHDPTRDQSWYVNNKLRQR 2256
QY 602 LREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFSLPETSALSQAQLCHQGA 661
Db 2257 LLEEYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESPHLTQELR 2316
QY 662 SLPP-----DHRMLYAQMDRAVFAQVKAAGCALQ 690
Db 2317 LLKEEINYDDKL---QVKNILYHAVKEMVRALK 2346

RESULT 12

US-10-193-874-16
; Sequence 16, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-16

Query Match 19.7%; Score 728.5; DB 4; Length 2540;
Best Local Similarity 29.1%; Pred. No. 5.4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLSGLGDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHGLFNTHWRC 120
Db 1802 SKYILDIIIGDKFCQLVTSEKTSALSVYKDAK-IWKRAVRGVREMDACEATLFTNIHWVC 1860

QY 121 SHCSHRLCVACRIAGAKGNREKTSQEQHT--DDCAQEAGHAACSLILTOFVSSQALAE 179
Db 1861 QKQFVVCLDYK---AKERKSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1916
QY 180 STVMQAWAKFDIRGHCF-----QVDARVWAPDGGQ-----KEPTE 218
Db 1917 LDAMHTLREKYGIKSHCHCTNKQNLQVGNPPTMNGVSQVLQNLVNLHNSKISLCMPESQOQ 1976
QY 219 KTPPTPOPSCNGSDNRTKIKER---TPDSTESP----- 249
Db 1977 NTPPKSEK--NGSSPESDVTGDKLTPPESQSPHLADLABQKAREEKENKELTLEN 2034
QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLGHRHMAFAPV--- 289
Db 2035 QIKEREQDNSESPNGRTSPLVSQNNQEGSTLRDLTTTAGKLRVGSTDAGIAFAPVYSM 2094
QY 290 -TPALPSDDRIITNLDIIIAQVVERKI-----QBKALGPGLRAGSGLR 331
Db 2095 GAPSSKSGRTMPNILDIIASVVENKI PPSKTSKINVKPELKEEPESIIISAVDENNKLY 2154
QY 332 KGLSLPLSPVTRTLSPGALLMLQEPKPKGHFLFOEHRWQOPVLVSGIQKTLRLSLWG 391
Db 2155 S--DIPHSWICEK-----HILMKDYKNSNNWKLKFCWKQGOPAVVSGVHKOMNISLWK 2207
QY 392 MEALG--TLGGQVQSLTALGPPQPTNLDSTAFWEGFSHPETRPKLDGSLVLLHRTLGDKD 450
Db 2208 AESSISLDFGDHQAADLLNCKDSIISANVKEFMDGFEEVSKRQKNSGETVWLK-----KD 2263
QY 451 -----ASRVNQLVSSLPPEYCAHQKLNLASVPLGLTLHP--LEPOLMAAYGV 498
Db 2264 WPSGDFKTMPPARYEDLLKSLPLPEYCNPEGKFNLAHLP--GFFVRPDLGPRLC SAYGV 2322
QY 499 NSHRGH--LGTKNLCVSDLSILVHAEAQLPPWYRAQKDFLSGLDGEGLWS----- 549
Db 2323 VAAKHODIGTTLNHLIEVSDVNNILVYVGIAGKNGILSKAGILKPFEEEDDDILRKRLKD 2382
QY 550 -----PGSQSTVWHVFRADAQRIIRFLQWVC--PAGAGTL---EPGAPGSCYLDAGLRRR 601
Db 2383 SSEIFGA---LWHIYAGKDVKIREFLQKISKEQGLEVLPEHDPTRDQSWYVNNKLRQR 2438
QY 602 LREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFSLPETSALSQAQLCHQGA 661
Db 2439 LLEEYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTEDFVSPHELVESPHLTQELR 2498
QY 662 SLPP-----DHRMLYAQMDRAVFAQVKAAGCALQ 690
Db 2499 LLKEEINYDDKL---QVKNILYHAVKEMVRALK 2528

RESULT 13

US-10-193-874-15
; Sequence 15, Application US/10193874
; Publication No. US20030064396A1
; GENERAL INFORMATION:
; APPLICANT: Jenapharm GmbH & Co. KG
; TITLE OF INVENTION: Comodulators of Nuclear Receptors
; FILE REFERENCE: 52145
; CURRENT APPLICATION NUMBER: US/10/193,874
; CURRENT FILING DATE: 1998-11-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-193-874-15

Query Match 19.7%; Score 728.5; DB 4; Length 2552;
Best Local Similarity 29.1%; Pred. No. 5.4e-47;
Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
QY 61 AKHLSGLGDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCCSRCHGLFNTHWRC 120

Db 1814 SKYILDIIGDKFCQVLTSEKTAISLVKWKDAK-IANKRAVRGVREMCDAEATLNFHWWC 1872
 QY 121 SHCSHRLCVACGRIAGAGNREKTSQEOHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
 Db 1873 QKCGFVCLDYK----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1928
 QY 180 STVMHQAQAFDIRGHCFC----QVDARVWAPGQGGQ-----KEPTE 218
 Db 1929 LDAMHTLREKYGKSHCHCTKNQNLQVGNFPTMNGVSQVQLQNLVHNSKISLCMPESQ 1988
 QY 219 KPTPTPQPCNGSDNKTOKIEB---TPDSTESP----- 249
 Db 1989 NTPPKSEK--NGGSSPESDVTGNKLTTPESQPLHLADLAQKAREEKENKELTLEN 2046
 QY 250 -----AEDGAGR-SLPC-----PSLCELLASTAVKCLGHDRIHMAFAPV--- 289
 Db 2047 QIKEREQDNSESPNGRTSLVSONNEQSTLRLDTTTAGKLRVGTDAIAFAPVYSM 2106
 QY 290 -TPALPSDDRTNILDSTIAQVVERKI-----QKALGPGLRAGSLR 331
 Db 2107 GAPSSKSGRTMPNILDITIASVVENKIPPSKTSKINVKPELKEEPBESIISAVDENNKLY 2166
 QY 332 KGLSLPLSPVTRTSLPPGALLMLOEPRPKGPHLFOEHWQOGQVLSVGIQKTLRLSLWG 391
 Db 2167 S--DIPHSWICEK-----HILWLKDYKNSSNWKLFKECWQOGQPAVVSGVHKKNISLWK 2219
 QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFSPHETPRPKLDGSLVLLHRTLTGDKD 450
 Db 2220 AESISLDFGHDADLLNCKDSIISANVKEFDGFEVSKRQKSGETVVVLK----KD 2275
 QY 451 -----ASRVQLVSSLPPEYCAHOGKLNLSYLPGLTLHP-LEPOLMAAYGV 498
 Db 2276 WPSGEDFKTMMPARYEDLLKSLPEYCNPEGKFNLAHLPGFFVPRDLGRLCSAYGV 2334
 QY 499 NSHRGH-LGTYNLCVSVLSILVHAEALPMPYRAQKDFLSGLDGEGLWS----- 549
 Db 2335 VAAKDHIGTNTNLHIEVSDVNILVYVGIAGNGILSKAGILKKFEEEDDDILRKRLKD 2394
 QY 550 ----PGSOTSTVWHVFRQAQRIIRFLQVVC-PAGAGTL---EPGAPGSCYLDAGLRRR 601
 Db 2395 SSRIPGA-----LWHIYAGKVDKIREFLQKISKEQGLEVLPEHPIRDQSWYVYKLRQR 2450
 QY 602 LREWGVSCWTLQAQPAEVLVAPAGAPHQVGLVSTISVTOHFLSPETSALSALQCHQGA 661
 Db 2451 LLEBYGVRTCTLIQFLGDAIVLPAGALHQVNFHSCIQVTEDFVSPHFLVESFHLTOELR 2510
 QY 662 SLPP-----DHRMLYAQMDRAVQAVKAAVGAALQ 690
 Db 2511 LKKEEINYDDKL---QVKNILYHAYKEMVRALK 2540

RESULT 14
 US-10-719-993-589
 ; Sequence 589, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001496
 ; CURRENT APPLICATION NUMBER: US/10/719,993
 ; CURRENT FILING DATE: 2003-11-24
 ; NUMBER OF SEQ ID NOS: 55342
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 589
 ; LENGTH: 2210
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-719-993-589

Query Match 19.4%; Score 719; DB 5; Length 2210;
 Best Local Similarity 30.0%; Pred. No. 2.4e-46;
 Matches 207; Conservative 108; Mismatches 250; Indels 124; Gaps 23;

QY 61 AKHLLSGLDRLCRLLRKRERENLAWAQREGQGPAMTEDSPGIPHCSCRRCHGLENTHWRC 120
 Db 1515 SKYILDIIGDKFCQVLTSEKTAIPRVKDAK-IANKRAVRGVREMCDAEATLNFHWWC 1573
 QY 121 SHCSHRLCVACGRIAGAGNREKTSQEOHT-DDCAQEAAGHAACSLILTFQVSSQALAE 179
 Db 1574 QKCGFVCLDYK----AKERSSRDKELYAMKCVKGQPHDKHLMPTQIIPGSLVTL 1629
 QY 180 STVMHQAQAFDIRGHCFCQVDARVWA---PDGGOQKPEPTKPTTPOPSNGSDNRRTK 236
 Db 1630 LDAMHTLREKYGKSHCHCTKNQNLQVGNFPTMNGVSQSQQNTTPKSEK--NGSSSPES 1687
 QY 237 DIKEB---TPDSTESP-----AEDGAGR- 256
 Db 1688 DVTGNKLTTPESQPLHLADLAQKAREEKENKELTLENOIKEREQDNSESPNGRT 1747
 QY 257 SPLPC-----PSLCELLASTAVKCLGHDRIHMAFAPV---TPALPSDDRTNILDSTI 306
 Db 1748 SPLVSONNEQSTLRLDTTTAGKLRVGTDAIAFAPVYSMGAPSSKSGRTMPNILDIT 1807
 QY 307 IAQVVERKI-----QKALGPGLRAGSLRKLGLSLPLSPVTRTSLPPG 349
 Db 1808 IASVVENKIPPSKTSKINVKPELKEEPBESIISAVDENNKLYS--DIPHSWICEK----- 1860
 QY 350 ALLMLOEPRPKGPHLFOEHWQOGQVLSVGIQKTLRLSLWGMEALG-TLGGQVQSLTAL 408
 Db 1861 HILWLKDYKNSSNWKLFKECWQOGQPAVVSGVHKKNISLWKAESISLDFGHDADLLN 1920
 QY 409 GPQPTNLDSTAFWEGFSPHETPRPKLDGSLVLLHRTLTGDKD-----ASRVQL 457
 Db 1921 KOSIISANVKEFDGFEVSKRQKSGETVVVLK----KDWPSGEDFKTMMPARYEDL 1976
 QY 458 VSSLPPEYCAHOGKLNLSYLPGLTLHP-LEPOLMAAYGVNSHRGH-LGTKNLCVSV 515
 Db 1977 LKSLPLPEYCNPEGKFNLAHLPGFFVPRDLGRLCSAYGVVAAKDHIGTNTNLHIEVS 2035
 QY 516 DLISILVHAEALPMPYRAQKDFLSGLDGEGLWS-----PGSOTSTVWHVFR 563
 Db 2036 DVVNILVYVGIAGNGILSKAGILKKFEEEDDDILRKLRKOSSEIPGA---LWHIYAG 2091
 QY 564 QDAQRIIRFLQVVC-PAGAGTL---EPGAPGSCYLDAGLRRRLREWGVSCWTLQAQPA 619
 Db 2092 KVDKIREFLQKISKEQGLEVLPEHPIRDQSWYVYKLRQLRLEBYGVRTCTLIQFLGD 2151
 QY 620 AVLVPAGAPHQVGLVSTISVTOHFLSP 648
 Db 2152 AIVLPAGALHQVNFHSCIQVTEDFVSPE 2180

RESULT 15
 US-10-719-993-588
 ; Sequence 588, Application US/10719993
 ; Publication No. US20040265849A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001496
 ; CURRENT APPLICATION NUMBER: US/10/719,993
 ; CURRENT FILING DATE: 2003-11-24
 ; NUMBER OF SEQ ID NOS: 55342
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 588
 ; LENGTH: 2353
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-719-993-588

Query Match 19.3%; Score 714; DB 5; Length 2353;
 Best Local Similarity 29.1%; Pred. No. 6.6e-46;
 Matches 219; Conservative 117; Mismatches 263; Indels 154; Gaps 27;

The Fog Bank (1870)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:03:15 ; Search time 44 Seconds
(without alignments)
1515.414 Million cell updates/sec

Title: US-10-712-629B-18
Perfect score: 3704
Sequence: 1 VTQCSCVQAGEVGLTGH.....MDRAVFQAVKAAVGALEAK 693
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3682	99.4	1182	2 I48378	hairless protein -
2	604.5	16.3	1214	2 S28499	probable finger pr
3	338	9.1	851	2 D86254	hypothetical prote
4	337	9.1	906	2 T01440	hypothetical prote
5	246	6.6	730	2 B85013	hypothetical prote
6	246	6.6	730	2 T10539	hypothetical prote
7	240.5	6.5	1629	2 T06461	DNA-binding protei
8	237	6.4	1641	2 T10955	early modulin bind
9	232	6.3	1701	2 T43213	ENBPI protein - ba
10	205.5	5.5	728	2 T05151	hypothetical prote
11	181.5	4.9	950	2 F86222	hypothetical prote
12	146.5	4.0	2414	2 A54277	transcription adap
13	145.5	3.9	2440	2 S39162	transcription coac
14	145.5	3.9	2441	2 S39161	CREB-binding prote
15	127.5	3.4	1479	2 T42710	mannose receptor,
16	122.5	3.3	594	2 S33561	ref(2)P protein -
17	122	3.3	3938	2 T42761	Bassoon protein -
18	121	3.3	778	2 I38487	tastin - human
19	119	3.2	711	2 S05381	VGF8a protein prec
20	117.5	3.2	862	2 T42889	hypothetical prote
21	116	3.1	907	2 A24938	hypothetical T2 pr
22	115	3.1	761	2 S20458	pqqf protein - kle
23	114	3.1	442	2 JC4978	oxidative stress p
24	113.5	3.1	2509	2 G01880	fatty-acid synthas
25	112.5	3.0	728	2 S71467	diacylglycerol kin
26	112.5	3.0	739	2 A83015	primosomal protein
27	112.5	3.0	1616	2 I37183	gene APXL protein
28	111.5	3.0	3942	2 T42730	Bassoon protein -
29	111	3.0	915	2 S36327	clathrin assembly

30 111 3.0 1170 2 A57650 repair protein xpg
31 110.5 3.0 572 2 T37128 hypothetical prote
32 110 3.0 1072 2 A37127 microtubule-associ
33 110 3.0 4861 2 S71752 giant protein p619
34 109.5 3.0 1356 2 A45445 janusin precursor,
35 108.5 2.9 1113 2 T00271 hypothetical prote
36 108 2.9 1123 1 WMBEH7 UL37 protein - hum
37 107.5 2.9 1673 2 I50806 complement compone
38 107.5 2.9 5149 2 F83345 probable non-ribos
39 107 2.9 892 2 B46203 mating type A alph
40 107 2.9 2044 2 T13704 still life protein
41 107 2.9 2133 2 T42763 coagulation factor
42 106.5 2.9 353 2 B72468 hypothetical prote
43 106.5 2.9 427 2 JN0785 carbon catabolite
44 106.5 2.9 896 2 S36326 clathrin assembly
45 106.5 2.9 1045 2 JC5795 CDEP protein - hum

ALIGNMENTS

RESULT 1

I48378
hairless protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48378
R;Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Best, S.; Stoye, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994
A;Title: Structure and expression of the hairless gene of mice.
A;Reference number: I48378; MUID:94329587; PMID:8052649
A;Accession: I48378
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1182 <RES>
A;Cross-references: UNIPROT:Q61645; UNIPARC:UPI0000028C24; EMBL:Z32675; NID:G531706; PI

Query Match 99.4%; Score 3682; DB 2; Length 1182;
Best Local Similarity 99.4%; Pred. No. 3.3e-257;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VTQCSCVQAGEVGLTGHGSKRRSPLEERKQLEEDSSATSEEGGGPGPEASLNKGL 60
DB 490 VTQCSCVQAGEVGLTGHGSKRRSPLEERKQLEEDSSATSEEGGGPGPEASLNKGL 549
QY 61 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
DB 550 AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 609
QY 121 SHCSHRLCVACGRIAGAGKNRKETSQEQHTDDCAQEAAGAACSLTLTQFVSSQALAE 180
DB 610 SHCSHRLCVACGRIAGAGKNRKETSQEQHTDDCAQEAAGAACSLTLTQFVSSQALAE 669
QY 181 TVMHWAKFDIRGHCFQVQDARVWAPGDGGQKQKPTPTPTPQSCNGDSNRTKDKE 240
DB 670 TVMHWAKFDIRGHCFQVQDARVWAPGDGGQKQKPTPTPTPQSCNGDSNRTKDKE 729
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKULGHDRIHMAFVTPALPSDDRIT 300
DB 730 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKULGHDRIHMAFVTPALPSDDRIT 789
QY 301 NILDSTIIAQVVRKIQEALGRLRGLSLPLSPVTRLSPPGALLWLQPRPK 360
DB 790 NILDSTIIAQVVRKIQEALGRLRGLRGLSLPLSPVTRLSPPGALLWLQPRPK 849
QY 361 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQSLTALGPQPTNLDSTA 420
DB 850 HGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPQPTNLDSTA 909
QY 421 FWEHGFHSHETRPKLDGEGSVLLHRLTGDKASRVQNLVSSLPPEYCAHQGKLNASYLP 480
DB 910 FWEHGFHSHETRPKLDGEGSVLLHRLTGDKASRVQNLVSSLPPEYCAHQGKLNASYLP 969

QY 481 LGTLHPLPQWAAVGVNSHRHGLTKNLQVEVSDLSILVHAEAQLPWYRAQKDFLS 540
Db 970 LGTLHPLPQWAAVGVNSHRHGLTKNLQVEVSDLSILVHAEAQLPWYRAQKDFLS 1029
QY 541 GLDGEGLWSPGTSVWVHVFRAQDAQRIRRFLOWKVPAGAGTLEPGAGSCVYLDAGLR 600
Db 1030 GLDGEGLWSPGTSVWVHVFRAQDAQRIRRFLOWKVPAGAGTLEPGAGSCVYLDAGLR 1089
QY 601 RLREEVGSCWTLLOAPGAVLVFAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQ 660
Db 1090 RLREEVGSCWTLLOAPGAVLVFAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQ 1149
QY 661 ASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 693
Db 1150 ASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 1182
RESULT 2
S28499
probable finger protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C:Accession: S28499
R;Hoog, C.; Schalling, M.; Grunder-Brundell, E.; Daneholt, B.
submitted to the EMBL data Library, June 1991
A:Description: Analysis of a murine germ cell-specific transcript that encodes a putative
A:Reference number: S28499
A:Accession: S28499
A:Molecule type: mRNA
A:Residues: 1-1214 <HOO>
A:Cross-references: UNIPROT:Q63679; UNIPARC:UPI000013774D; EMBL:X59993; NID:g57503; PIDN
A:Experimental source: strain Sprague Dawley
C:Keywords: DNA binding; zinc finger
Query Match 16.3%; Score 604.5; DB 2; Length 1214;
Best Local Similarity 26.7%; Pred. No. 2.5e-35;
Matches 195; Conservative 114; Mismatches 289; Indels 131; Gaps 23;
QY 61 AKHLLSGLRCLRLKKEALWAQREGQGPAMTSDSPGPHCCSRCHGLFNTHWRC 120
Db 502 AKILANIGDFHFCQVISEKAMSTIEPHRQ-VANKEAVKGVREMCDDVTIFNLHWVC 560
QY 121 SHCSHRLCVACGRIAGKAKREKTGSGEHT-DDCAQEAAGAACSLILTOFVSSQALAE 179
Db 561 PRCGFGVCVDCYRL--KRKNCQGAAYKTFSWIRCVKSKIHEPENLMPQTIIIPGKALYDV 618
QY 180 STVHQWAKFDIRGHCF--QVDARVWAPDGGQOKEPTEKPTTPQPSGNGDSNR 234
Db 619 GDVHSVRKAWGKANCPCSNROPKLFSKPALKEDLKQASLSGK-----PSLGTWVQQ 672
QY 235 TKDKEETPDSTESPAEDGAGRSLPCPS-----LCELLASTAVKLCGLGHDRIHMAFA 287
Db 673 SSPVLEPAAVCGEAPSPASNVKPI-CPANTSPLNLWADLITSGNVK-----ENKEKOLTM 727
QY 288 PV-----TPALP-----SDRLTN- 301
Db 728 PILKNEIKPLPLPLNKSSTVLHTFNSTILTPVNNNSGFLRNLNLSSTGKTEGLKNT 787
QY 302 --TILDSIIAQVVERKIOEKALGPLRAGSL-RKGLSLPLSPVTRLSPP-----CALL 352
Db 788 PKLDDDFASLVQNK-----TSSDLSKRPOGLTIKPSILGFTTHYMLCDNRLL 836
QY 353 WLQEPKPKHGLPHFOEHWROGQPVLVSGIOKTLRLSLWGMEALGTGG--QVQSLTAGP 410
Db 837 CLQDPNNKSNVNFRECKQGPVWVSGVHHKLNTELWKFPESFRKEGEQEVDLVNCRTN 896
QY 411 PQPTNLDSIAFWEGFHPETRPKLD---EGSVLLH-----RTLGDKDAQRVQNLVSSLP 462
Db 897 EIIITGATVDFDGFEDFVNRLKNEKEKEPMLVKLKDWPPEGEDFRDWMPSRFDLMANIP 956
QY 463 LPEYCAHQKLNLSYPLGLTLHPLPQWAAVGVNSHRHGLTKNLQVEVSDLSIL 521
Db 957 LPEYTRRDGKLNLSRFLNPFVVRPDLGPKMYNAYGLITTPEDRKYGTTLNLHLDVSDAANVM 1016

QY 522 VHAEAQLPWYRAQKDFLSGLDG-----EGLMSPGQSQTSTVHVHVFRAQDAQRIR 570
Db 1017 VVVGIPKQCEQEVEVLRTIQDGSDELTIKRFIEGKEKPGA-----LHVIYAAKDKTEKIR 1072
QY 571 RLQWVC-----PAGAGTLEPGAGSCYLDAGLRRLREBWGVCWTLLOAPGSAVLV 623
Db 1073 EFLKKVSEEQGENPADH---DPIHQSWYLDRLSRKRLYQBYGQWAIQVFLGDDVVF 1129
QY 624 PAGAPHQVQGLVSTISVTHFLSPETSALSQAQCHQASLPPDH--RMLYAQMDRAVFOA 681
Db 1130 PAGAPHQVHNLSCIKVAEDFVSPHVKHCFWLQTEFRHLSTQHTTHNEDKLVQNVYIHA 1189
QY 682 VKAAVQALQ 690
Db 1190 VKDAVAMLK 1198
RESULT 3
D86254
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D86254
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurob, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-851 <STO>
A:Cross-references: UNIPROT:O65384; UNIPARC:UPI00000A00D6; GB:AE005172; NID:g3157933; P
C:Genetics:
A:Map position: 1
Query Match 9.1%; Score 338; DB 2; Length 851;
Best Local Similarity 22.2%; Pred. No. 2.5e-16;
Matches 148; Conservative 91; Mismatches 192; Indels 236; Gaps 30;
QY 106 CSRCHGLFNTHWRCSHCSHRLCV-ACGRIAG-----AGKREKTGSGEHTDD----- 153
Db 322 CNHCATSIVDLHRSCPKCYELCNCCQEIRGEPSSSVSEDTKTPSIKWNADENGSI 381
QY 154 CA-QEAGHAACSLILTOFVSSQALAEIETVMHQWAKFDIRGHCFQCDVARDVWAPDGGQ 212
Db 382 CAPKELGGCDSVL-----ELKRILPVTWMS-DL-----E 410
QY 213 QKEPT-----EKTPTTPQPSGNGDSNRDKIETPDSTESPAEDGAGRSLPCPSLCEL 267
Db 411 QKAETFLASYSIKPPMSYCRCSDDMSMK-----RKAASRDGSSDNYLSPDSDLV 461
QY 268 LASTAVKLCGLGHDRIHMAFAPVTPALPSDDRIITNILDSTIAQVVERKIOEKALGPLRAG 327
Db 462 LK-----QELL----- 468
QY 328 SGLRKGLSLPLSPVTRLSPPGALLWLQEPKPKHGLPHFOEHWROGQPVLV-SGIOKTLR 386
Db 469 -----H-FQEHWSKGEPIVTRNALNNTAG 491
QY 387 LS-----LWGMEAL-----GTGGQVQSLTAGLPQPTNLDSIAFWEGFHPETR----P 432
Db 492 LSWEPVWVW--RALCENVDSAISSNMSDVKAIDCLANCEINTLCFFEGYSGKRTYENFWP 549
QY 433 KLDEGSVLLHRLHRTLGDKD--ASRVQNL-----VSSLPLPEYC-AHQGKLNLSYLP 480

585	SKGRTYENFWPEMLKLDKWPSPDFEDILL-----PRHDEFISALPFOEYSDPRTG	635
472	KLNLASVPLGLTLHPLEBPOLWAAVGNVSH--RGHLGTKNLCVEYSDLISILVH-AEAQL	528
636	ILNLATKLPEGFIPDLGPKTYIAYGIPDELGRGDSVTKLHC-DMSDAVNILTHTAEVTL	694
529	PPWYRAQKDFLSGL-----DGEGLWSP-----GSQT-STV	557
695	S--QEQISSVKALKQKHKLQNKVDKQSTDCCNKEKEEELNNPESISSENEETGSAL	751
558	WHVFRADAQRIPLFLOVPCAGAGT-----LEP	586
752	WDIFRREDVPKLBYLRKHCKEFRHTYCSPTVTKTPTCINPMTNLPVLTVSSFOVYHP	811
587	GAPGSCYLDAGLRRLREEMGVSVCWTLLQAQGEAVLVPAGAPHQVQGLVSTISVTHFLS	646
812	IHDOSCVLTLEHKRKLKAEYGIPEWTFVQKLGAEVFIAGCPHQVRNLKSCCTKVAVDVFS	871
647	PETSALSAQLCHOGASLPDDHR	668
872	PENIHECLRLEETEEFROLPKPKHK	893

RESULT 5

```

RESULT 5
B85013
  hypothetical protein AT4g00990 [imported] - Arabidopsis thaliana
  C/Species: Arabidopsis thaliana (mouse-ear cress)
  C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change
  C/Accession: B85013
  R/anonymous: The European Union Arabidopsis Genome Sequencing Co
  Nature 402, 769-777, 1999
  A/Title: Sequence and analysis of chromosome 4 of the plant Arab
  A/Reference number: B85001; MUID:20083488; PMID:10617198
  A/Accession: B85013
  A/Status: preliminary
  A/Molecule type: DNA
  A/Residues: 1-730 <STO>
  A/Cross-references: UNIPROT:Q9SV29; UNIPARC:UPI00000AB5E5; GB:NC
  C/Genetics:
  A/Gene: AT4g00990
  A/Map position: 4

```

Query Match	6.6%; Score 246; DB 2; Length 730;
Best Local Similarity	19.3%; Pred. No. 9e-10;
Matches	154; Conservative 91; Mismatches 236; Indels 318; Gaps 30
Qy	91 QGPWNTSDSGIPIHC-----CSRCHGLFNTHWEC--SHCSHRLCVACGRIAGAK 139
Db	43 RGHEVTE--ANIKRKLDPSRIYCDLRTCRTSIANPHRSCPNKNCSDVICLSCCKELSEGF 1000
Qy	140 NREKTGSO--EQTHTDDCAQEAAGHAACSLILTFQVSSQALAEELSTVMHQAWAKFDIRGHCF 197
Db	101 HQERDGKKNAGKGYEGRIPAGQCK----- 125
Qy	198 CQVDARYVWPGDGGQOQKEPTKPTPTPOPSCNGDSNRKTDKIETPDSTESPAEDGAGRS 257
Db	126 --DSDAYVPLHFSTWKLNSDSSIPCPKCEGCGGTSTLELRR----- 165
Qy	258 PLPCPSCELLASTAVKLCGLH-----DRIHMAFAPVTPALPDDIRTNILDSIIAQVVE 312
Db	166 -LWKRDVVEKLITNAEKCCTLNFRPTDVIDVHECSCSCTNS-----DSIRRQAAF 213
Qy	313 RKIOEKALGPCLRAGSLRGKLSPLSPVRTLSPPGALLWLQEPKPKGHGFLHQEHWRQ 372
Db	214 RKNANDNF-----LYSPNA-VDLAEDDIAH-----FQFHMWK 244
Qy	373 GQPVLVSGI-OKTLRLS-----LM-----GMEALGTGQVOSLTALGPPOPTN 415
Db	245 AEPVIVRNVLKTSGLSWEPVWMVWRACREMDPKRGTTEEYT---KVALDCLDWCE--VE 300
Qy	416 LDSTAFWEGF-----SHPETRPKLDGQSVLLHRLTGDGKASRVQNLVSSLPLPYVC 467

Db 301 INLHOFEGYLGGRMHKNGWPEMLKWKOWPPSDLFEKRLPRHNA-----EFIAALPPFDYT 356
QY 468 -AHQGLNLASVLPGLTLHLPLEPQWAAAYGVNS--HRGHLGTKNLCVSVDSILSVH- 523
Db 357 DPKSGILNLATRPFGESLKPDLGPKTYIAYGPHFELNRGDSVTKLHC-DISDAVNLVTH 415
QY 524 AEAQLPP-----WYRAQKFLSLDGEGL----- 547
Db 416 AKVEIPPVKYQNIKVHOKKYAEAMLKQOYSGQVKEASELENKSMKEVDESKKLDKAA 475
QY 548 ----- 547
Db 476 NEEQNSNSRPSGSGAEKVIISKIARIELSHSVYVYKMLLNWENGLMPTLLATPPC 535
QY 548 -----WSPGSOTS-----TVHVFRAQDAQRIIRF 572
Db 536 DTEDNTPQAVSTSVESIQEQLDAPKETDGTNTNERSKAVHGGAVWDIFREDVPKLIQF 595
QY 573 LQ-----MVCAPAGATLEPGAGCYLDAGLRRLRREBNGVSWTLLQA 616
Db 596 LKRHEHEFRHFNNEPLESVIHPHDQTM-----FLSDSKKQLKEBFDIEPMTFEQH 647
QY 617 PGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSQALCHOGASIPDHRMLYAQMDR 676
Db 648 LGEAVFIPAGCPHQVNR---QVALDFVAPESVECLRLTQEFRRLPKDHS---SSEDK 700
QY 677 AVFQ--AVKAAVGALQEA 693
Db 701 LELKKIALYAASSAIREVK 719
RESULT 6
T10539
hypoetical protein F313.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10539
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17080
A:Accession: T10539
A:Molecule type: DNA
A:Residues: 1-730 <BEV>
A:Cross-references: UNIPROT:Q9SV29; UNIPARC:UPI000000AB5E5; EMBL:AL080237; GSPDB:GN000062;
A:Experimental source: cultivar Columbia; BAC clone F313
C:Genetics:
A:Gene: ATSP:F313.10
A:Map position: 4
A>Note: intron positions not resolved
Query Match 6.6%; Score 246; DB 2; Length 730;
Best Local Similarity 19.3%; Pred. No. 9e-10;
Matches 154; Conservative 91; Mismatches 236; Indels 318; Gaps 30;
QY 91 QGPAMTEDSPGIPHC-----CSCHHGLFNTHWRC---SHCSHRLCVACGRIAGAK 139
Db 43 RGHVTE--ANIKRCLDPSEIYCDLCRTSIANFHRSCPNKNGCSVDICLSCKELSEGF 100
QY 140 NREKTSQ--EQHTDDCAQAGHAACSLILTFQVSSQALAELSVVMQAWKPDIRHCF 197
Db 101 HQBRDGKNAEGKGYECRIAPAGQK----- 125
QY 198 CQVDARVWAPGDGQOQKEPTKPTTPQPSQNGDSNRTKDIKEETPDSTESPAEDGAGRS 257
Db 126 ---DSDAYVPLHFTWKLNSDSSIPCPPEKCGGCGTSTLELR----- 165
QY 258 PLCPSPCELLASTAVKLCIGH-----DRIHMAFAPVTPALPSPDDRITNILDIIAQQVE 312
Db 166 -LWKRDVWEKLIITNAEKCITLNFRTDVIDIHECSCSCTNS-----DSIRQAAF 213
QY 313 RKIQEKALGFLRAGSLRGLSLPLSPVTRTLSPGALLWLQPRPKHGPHLFOEHWQ 372
Db 214 RKNADNF-----LYSPNA-VDLAEDDIAH----FQFHWK 244

QY 373 GQPVLVSGI-OKTLRLS-----LW-----GMEALGTGQVQSLTALGPPQPTN 415
Db 245 ABPVTVRVNLEKTSGLSWPEVMWRACREMDPKRGTEET---KYKALDCLDWCE-VE 300
QY 416 LDSTAFWEGF-----SHPETRPKLDEGSVLLHLHRTLGDKDSRVQVNLVSLPLPEYC 467
Db 301 INLHOFEGYLGGRMHKNGWPEMLKWKOWPPSDLFEKRLPRHNA-----EFIAALPPFDYT 356
QY 468 -AHQGLNLASVLPGLTLHLPLEPQWAAAYGVNS--HRGHLGTKNLCVSVDSILSVH- 523
Db 357 DPKSGILNLATRPFGESLKPDLGPKTYIAYGPHFELNRGDSVTKLHC-DISDAVNLVTH 415
QY 524 AEAQLPP-----WYRAQKFLSLDGEGL----- 547
Db 416 AKVEIPPVKYQNIKVHOKKYAEAMLKQOYSGQVKEASELENKSMKEVDESKKLDKAA 475
QY 548 ----- 547
Db 476 NEEQNSNSRPSGSGAEKVIISKIARIELSHSVYVYKMLLNWENGLMPTLLATPPC 535
QY 548 -----WSPGSOTS-----TVHVFRAQDAQRIIRF 572
Db 536 DTEDNTPQAVSTSVESIQEQLDAPKETDGTNTNERSKAVHGGAVWDIFREDVPKLIQF 595
QY 573 LQ-----MVCAPAGATLEPGAGCYLDAGLRRLRREBNGVSWTLLQA 616
Db 596 LKRHEHEFRHFNNEPLESVIHPHDQTM-----FLSDSKKQLKEBFDIEPMTFEQH 647
QY 617 PGEAVLPAGAPHOVQGLVSTISVTOHFLSPETSALSQALCHOGASIPDHRMLYAQMDR 676
Db 648 LGEAVFIPAGCPHQVNR---QVALDFVAPESVECLRLTQEFRRLPKDHS---SSEDK 700
QY 677 AVFQ--AVKAAVGALQEA 693
Db 701 LELKKIALYAASSAIREVK 719
RESULT 7
T06461
DNA-binding protein P03, chloroplast - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06461
R:Sato, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z15695
A:Accession: T06461
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1629 <SAT>
A:Cross-references: UNIPROT:O04698; UNIPARC:UPI000009FC7F; EMBL:X98744; NID:e995233; PI:
A:Experimental source: cv. Alaska
C:Genetics:
A:Genome: nuclear
C:Keywords: chloroplast
Query Match 6.5%; Score 240.5; DB 2; Length 1629;
Best Local Similarity 19.1%; Pred. No. 6.2e-09;
Matches 157; Conservative 63; Mismatches 221; Indels 383; Gaps 27;
QY 106 CSCHHGLFNTHWRC--HCSHRLCVACG---RIAGKGNREKTSQEQHTDDCAQAGH 160
Db 934 CQNCNTSIVNFHRSCPNKNCQVDCLTCTCTELRIGVHCIDIPASGNEE----- 981
QY 161 AACSLILTFQVSSQALAELSVVMQAWKPDIRHCFQVDARVWAPGDGQOQKEPTK 220
Db 982 -----WYDA-----PPEGI 990
QY 221 PPTPQPSQNGDSNRTKDIKEETPDSTESP--ABDGAGRSPLPCPSICELLASTAVKLCIG 278
Db 991 P-----WRAETNGSI PCPPKARGGGGCIATLSLRLLFE--ANWIDKTRG 1032

Query Match 6.3%; Score 232; DB 2; Length 1701;
Best Local Similarity 19.6%; Pred. No. 2.7e-08;
Matches 171; Conservative 71; Mismatches 259; Indels 370; Gaps 33;
QY 71 RLCRLRKEREALAWAQREQ-----GPAETDSFGIPHC-----CSRCHHG 112
Db 947 KLFYLLKTKLPLLOHTQREQKSELEVASIHGSLMVEEKDILQAQVDDDDRVVCDNCNTS 1006
QY 113 LFETHWRC--SHGSHRLCVACGRIAGAKNREKTGSEQHTDDCAQEAAGAACSLILTOF 170
Db 1007 IVNFHRSVNPYCRYDLCLTC-----PASGG--NEEMVNTPP----- 1029
QY 171 VSSQALAEISTVMQAWAKFDIGHCFQVDARWAPGDGGOQKEPTEKTPPTPQPSGNG 230
Db 1030 -----ELRNGVHSHK-----DI-----PASGG--NEEMVNTPP----- 1054
QY 231 DSNRTDKIETPDSTESP--AEDGAGRSLPCPSLCELL--ASTAVKLCIGHDRIHMAFA 287
Db 1055 ---ETIAWRAETNGSIPCPKARGCGGTATL---SURLFKANWIEKLTDRABELTIKYQ 1108
QY 288 PVTPALPSSDDRIITNILDSTIAQVVERKIOBKALGPGRLAGSLRKGSLPLSPVTRLSP 347
Db 1109 P-----PIVDLSLECESECSFEEDA-----AHNSARKAASRETGHDNLLYCP 1150
QY 348 PGALLMLQEPKPKHGHFLFOEHRQGPVLVSGIQK-----TLRLSLWME 393
Db 1151 DAIEIGDTE-----FDHFQHWIRGEPVIVRNYYKKGSLSDPMMVMAFRL----- 1198
QY 394 ALCTLGQVOSLTALGPPQ--PTNLDSTAFWEGF-----SHPETRPLKDEGSVLLH 443
Db 1199 AKNILDEADTFKAIACLDCWCEVOVNAFQFFKGLTGRRYRNQWPEMLKLDMPNPFPE 1258
QY 444 RTLGDKDASRVQNLVSLPLPEYC-AHQGKLNILASYLPLGLTLHP--LEPQLWAAAGV--N 499
Db 1259 DCLPRHCA---EFTAMLPSDSYTHPKSGILNLATKLP--TVLKPDLPKTYIAYGALBE 1312
QY 500 SHRGHGLTKNLCEVSDLSILVH-AEAQLPPWY----- 532
Db 1313 LSRGDSVTKLHC-DISDAVNILHTADVKTPAWQSKIHKKKYEVEDMRELYGLDSKA 1371
QY 533 ---RAOK----- 536
Db 1372 AGSRGRKRKRVRGVTVVDLKISEKEDINGRDSLTLLSQEKEDKLDREACVQEFSESTYSK 1431
QY 537 ----- 536
Db 1432 LDLNVSNQEVIDSPRFQFDNLSDSNFLVPRNDCESMLYDNVEQRCSPRDGCKGNTS 1491
QY 537 -----DFLSGLDGEGLWSPGSQT-----ST 556
Db 1492 VIDNQPCGGTKETTFVNGLSDSDISSSDIETDKIESVENEMPSNNLCGNDVHLETQYGSA 1551
QY 557 VVHVFRAQDAQRI-----RPLQWVCPAGAGTLEPGAGSCYLDAGLRRRLREEMGV 608
Db 1552 VMDIFRRQDVPKLEYLUNKHREPRHITSIPVNFVHPIHDQHFYLNKHKHKKQLKLEYGV 1611
QY 609 SCWTLQAPGEAVLPVAPAGHQVGLVSTTSVTHFLSPETSALSALCHQAGSLPPDHR 668
Db 1612 EPWTFEQLGAEVFIAPAGCHQVNRKPCIKVAMDVFVSPNVNCEVLTETEFRLLPKYHR 1671
QY 669 -----MLYQMDRAVFOAQAAGA 688
Db 1672 SKEDKLEIKMALYA-ADVAIAEATK-LVGA 1700

RESULT 10
T05151
hypotheical protein F18E5.50 - Arabidopsis thaliana
N;Alternate names: hypotheical protein T6K22.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05151; T05185
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; New

submitted to the Protein Sequence Database, August 1998
A;Reference number: Z15400
A;Accession: T05151
A;Molecule type: DNA
A;Residues: 1-728 <BEV>
A;Cross-references: UNIPROT:O65408; UNIPARC:UPI00000A0A08; EMBL:AL022603
A;Experimental source: cultivar Columbia; BAC clone F18E5
A;Accession: T05185
A;Molecule type: DNA
A;Residues: 1-728 <BEW>
A;Cross-references: UNIPARC:UPI00000A0A08; EMBL:AL031187
A;Experimental source: cultivar Columbia; BAC clone T6K22
C;Genetics:
A;Map position: 4
A;Introns: 208/3; 253/1; 298/1; 361/3; 471/3; 603/3; 628/1; 656/3; 696/3
A;Note: F18E5.50; T6K22.160

Query Match 5.5%; Score 205.5; DB 2; Length 728;
Best Local Similarity 19.8%; Pred. No. 7.4e-07;
Matches 121; Conservative 88; Mismatches 190; Indels 211; Gaps 25;
QY 220 TPPTPQPCNGSDSNRTDKIETPDSTESPADGAGRSPL-----PCPSLCE 266
Db 165 SPSP-----TTSNVSPCDVKVGEELPSMIKRRFRSKNIEPLP-IGK 206
QY 267 LLASTAVKLCIGHDRI-HMAFA-----PVTPALPSSDDRIITNILDSTIAQVVERKIOBKAL 320
Db 207 MODSQSVRSDI--DRVLHLHYAVCMLLPVLKEINAEHKVE-----VENDAERKVDLQRM 258
QY 321 GFLCRAGSLR-----KGLSLPLSPVR--TRLSPGALLWLOEPRPKHGHFLFOEH 369
Db 259 --CTRSSSVLRNLNSDDQSOESLSRKVGSKVCSNGIKSPKVLDDFHNNLEH---FQTH 312
QY 370 WRGQGVLSGTOKTLRLSLWME----- 396
Db 313 WSKGHFVIVRSVSKSSSLNWDVPVLFCHYLMNRNKTGTTDCMDWFEVIGVKQFFLG 372
QY 397 TLGGQVQSITALGPPPOPTNLDSTAFWEGFSPHETPRKPLDEGSVLLHRTLGDKDSRVQN 456
Db 373 SLRGAETNTC-----QERLKL-EG--WLSSSLFKEQFPNHYAE 408
QY 457 LVSSSLPLPEYC-AHQGKLNILASYLPLGLTLHPLEPQLWAAAGVNSHRGHL-----GTKNL 510
Db 409 ILNILPISHYMDPKRGLNLIAANLP--DTVQP--PDFGCLNLSYRSGBEYAPDSVKKL 464
QY 511 CVESVDSILSIHV-----AEALPPWYRAQK-----D 537
Db 465 GFETCDMDVILLVTTPTVSTNQICRKLKMKNGVRSKNPAKGRESRFDKGRDRDL 524
QY 538 FLUGLDE-----GLWSPGSQ-----TSTVHVFRAQDAQRI 570
Db 525 DYSSSDSESSQHCAGKRCGSEPEGEERESCNYSCEESLNTYGAQMDVFOQDVSKLL 584
QY 571 RF-----LQWVCPAGAGTLEPGAGSCYLDAGLRRRLREEMGV 623
Db 585 EYTKNSHLESMDSSKKVSHPLLEQSYLDYHFKARLKEEFDPWSDQCQGEAVIL 644
QY 624 PAGAPHQVGLVSTTSVTHFLSPETSALSALCHQAGSLPPDHRMLYQMDRAVFOAVK 683
Db 645 PACCPQIRKNKSCVNAVLLKFLSPHEVSESIKRVKELNQLP-----QSVK 689
QY 684 AAVGALQEAQ 693
Db 690 SKANKIEVKK 699

RESULT 11
F86222
hypotheical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: F86222
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Venter, J.C.; Davis, R.W.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86222
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-950 <STO>
A:Cross-references: UNIPROT:O04024; UNIPARC:UP100000A98940; GB:AE005172; NID:G2342679; PT:1
C:Genetics:
A:Map position: 1

Query Match 4.9%; Score 181.5; DB 2; Length 950;
Best Local Similarity 21.1%; Pred. No. 5.6e-05;
Matches 128; Conservative 76; Mismatches 208; Indels 195; Gaps 28

[illegible]

RESULT 12
A54277
transcription adaptor protein p300 - human
CfSpecies: Homo sapiens (man)
C.Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C.Accession: A54277: S60344

R; Eckner, R.; Ewan, M.E.; Newsome, D.; Gerdes, M.; Decaprio, J.A.; Lawrence, J.B.; Livi
Genes Dev. 8, 865-884, 1994
A; Title: Molecular cloning and functional analysis of the adenovirus E1A-associated 300
A; Reference number: A54277; PMID:95011587; PMID:7523245
A; Accession: A54277
A; Status: preliminary
A; Molecule type: mRNA
A; Cross-references: 1-2414 <ECK>
A; Note: in the authors' translation 941-Ser is shown after 961 and consequently, residu
R; Lundblad, J.R.; Kwok, R.P.S.; Laurence, M.E.; Harter, M.L.; Goodman, R.H.

Query Match 4.0%; Score 146.5; DB 2; Length 2414;
Best Local Similarity 19.0%; Pred. No. 0.061;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;

QY	29	LEE--KQLEEESSATSEGGCGPGE-----ASLNKGLAK---	62
Db	1513	LEESIKELQEBEEREKREENTNESPDVKGDSKNAKKXNKTSKNSLSRGNKKKPG	1572
QY	63	--HLLSGLDRLCLRLKEREALAWAQEGQPA-----MTEDSPGIP-----	103
Db	1573	MPNVSNDLSQKYATMEKHKEVP-FVIRLIAGPAANSLPPIVDPPDLIPCOLDGDRDAFL	1631
QY	104	-----HCCSRCHGLGFNTHWRCSHC--SHR	126
Db	1632	TLARDKHLEFFSLRRAQWSTMCMVLVELHTQSODRFVYTCNECKHHV-ETRWHTCVCEYD	1690
QY	127	LCVACGRIGAGKNREKTG-----SQEHTDDCAQEAHAA-----	162
Db	1691	LCITCYNTNKHDKMEKLGLDDESNNQAAATSPGDSRRLSIQRCIQSLVHACOCRN	1750
QY	163	--CSLLITQFVSSQALAEELSTMHQAWAKFDIRGHC-----FCQVD-----	201
Db	1751	ANCSL-----PSCQMKRV--VOHTKGCRKTNNGGCPICKQIALCCYHAKHQENKCPV	1803
QY	202	-----ARYMAPDGGQOK--EPTKEKTPPTP--QP	236
Db	1804	PFLCNIKQKLQQQLQHRLLQQAQLRRRWASMQRTGVVGGQGLPSPTPATPTPTGQOP	1863
QY	227	SCNGDSNRKYDKIETPTDSTESPAEDGAGRSPLPCPSCELLASTAVKULCLGHDRHMAF	286
Db	1864	T-----TPQTPQTSQP-----QPTPPNSM-----	1883
QY	287	APVTALPSPDDITNILDIIAQVVERKIQEALGPGRAGSLRGSLPLSPVPTRLIS	346
Db	1884	---PPVLP-----RTQAAGP---VSQKAAQVTPPTPPQTAQP	1916
QY	347	P-PGALLMLQERPP---KHGFHLFOEHWRRQGPVLVSGIQTKLRLSLGWMEALGTUGQOV	402
Db	1917	PLPG-----PPPTAVENAMQIQRAAETQRQWAVHQIFQRPIQHQPMPMTWPAKPG---	1966
QY	403	QSLTALGPPQPTNLDSTAFWEGF--SHPSTRPKLDEGVSLLHRTLGDKAARVQNLVSS	460
Db	1967	-----MNPPTMRGRSGHLEPGMGTPGMQQPPWSQGL-----PQPQQLQSG	2009

Search completed: November 25, 2005, 22:14:21
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 21:50:27 ; Search time 189 Seconds
(without alignments)
1611.056 Million cell updates/sec

Title: US-10-712-629b-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQBAK 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Genesep1980s.*
- 2: Genesep1990s.*
- 3: Genesep2000s.*
- 4: Genesep2001s.*
- 5: Genesep2002s.*
- 6: Genesep2003as.*
- 7: Genesep2003bs.*
- 8: Genesep2004s.*
- 9: Genesep2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3682	99.4	693	8	AD057501
2	3682	99.4	1182	5	AAE19798
3	3682	99.4	1182	6	ABU62544
4	3510	94.8	1207	5	AAE19797
5	3095.5	83.6	1189	6	ABU62543
6	3095.5	83.6	1189	5	AAE19796
7	3095.5	83.6	1189	6	ABU62542
8	3093.5	83.5	1189	2	AAV15218
9	3089.5	83.4	1189	2	AAV15217
10	3089.5	83.4	1189	6	ABG72775
11	3088.5	83.4	984	5	AAE19794
12	3088.5	83.4	984	6	ABU62545
13	3077.5	83.1	1189	5	AAE19795
14	3077.5	83.1	1189	6	ABU62541
15	3067.5	82.8	984	6	ABU62540
16	2759	74.5	1134	8	ADQ80370
17	729	19.7	1265	7	ADJ68989
18	728.5	19.7	1266	6	AAE19798
19	728.5	19.7	1473	9	AAE19798
20	728.5	19.7	2055	8	AD40509
21	700.5	18.9	1145	3	AAE19798
22	700.5	18.9	1417	7	AD40509
23	618.5	16.7	179	6	ABU70953
24	606.5	16.4	1236	7	ADD18797

25	606.5	16.4	1321	8	AD019904	Ado19904 Human PRO
26	606.5	16.4	1321	9	ADX06345	Adx06345 Cyclin-de
27	606.5	16.4	1337	4	AM78460	Am78460 Human pro
28	606.5	16.4	1338	4	AM79444	Am79444 Human pro
29	606	16.4	671	5	ABF41598	Abf41598 Human ova
30	604.5	16.3	1214	7	ADD46411	Add46411 Rat Prote
31	603.5	16.3	1212	5	ABF65070	Abf65070 Hypoxia-i
32	603.5	16.3	1213	7	ADD46413	Add46413 Human pro
33	577.5	15.6	511	4	AM41958	Am41958 Human pol
34	577.5	15.6	511	9	AEA21016	Aea21016 Novel hum
35	550.5	14.9	854	4	ABE63729	Abbe3729 Drosophil
36	448.5	12.1	1192	8	ADRI4639	Adri4639 Human NF-
37	362	9.8	875	8	AD062965	Ado62965 Transcrip
38	337	9.1	906	8	AD061723	Ado61723 Transcrip
39	309.5	8.4	917	8	AD062489	Ado62489 Transcrip
40	305.5	8.2	860	8	AD062492	Ado62492 Transcrip
41	305	8.2	963	9	ADWI7337	Adwi7337 Eucalyptu
42	297	8.0	747	8	ADX67858	Adx67858 Plant ful
43	297	8.0	1027	8	AD061957	Ado61957 Transcrip
44	279.5	7.5	363	8	ADX73095	Adx73095 Plant ful
45	277.5	7.5	953	9	ADW18304	Adw18304 Eucalyptu

ALIGNMENTS

RESULT 1

AD057501

ID AD057501 standard; protein; 693 AA.

XX AC AD057501;

XX DT 29-JUL-2004 (first entry)

XX DE Hairless protein interaction partner #35.

XX mouse; hairless protein; Hrt; ubiquitous receptor UR; MAP1A;
KW KIAA0930 protein; monocytes antigen CD14; sphingolipid activator protein;
KW beta-eynuclen; C11 protein; vesicle-associated membrane protein 2;
KW aldolase A; CIG-106 protein; hypothalamus protein HSMNP1; alpha enolase;
KW POM-2P3; quinone oxidoreductase; pumilio 1; VPS41; KIAA0614 protein;
KW splicing factor CCL4; ubiquitin; beta-mannosidase; hair growth.
XX Mus musculus.
XX US2004086945-A1.
XX 06-MAY-2004.
XX 02-JUN-2003; 2003US-00452858.
XX 03-JUN-2002; 2002US-0385414P.
XX (PROC) PROCTER & GAMBLE CO.
XX Sreekrishna K, Gerwe GS, Toerner DR;
XX WPI; 2004-430095/40.
XX N-PSDB; AD057500.
XX New composition comprising mouse truncated hairless protein-human
XX interacting partner protein or nucleic acid complexes, useful for
XX screening test compounds that inhibit or enhance hair growth.
XX Example 1; SEQ ID NO 93; 60pp; English.
XX The invention relates to a composition comprising a mouse truncated
XX hairless (Hrt) protein-human interacting partner protein or nucleic acid
XX complex. The human interacting partner protein comprises a molecule
XX selected from ubiquitous receptor UR, MAP1A, KIAA0930 protein, monocytes
XX antigen CD14, sphingolipid activator protein, beta-synuclein, C11
XX protein, vesicle-associated membrane protein 2, aldolase A, CIG-106
XX protein, hypothalamus protein HSMNP1, alpha enolase, POM-2P3, quinone

CC oxidoreductase, pumilio 1, VPS41, KIAA0614 protein, splicing factor
CC CCl.4, ubiquitin, beta-mannosidase. Also described are: methods of
CC assaying a test compound for agonist or antagonist activity for the above
CC composition; and methods of inhibiting or increasing hair growth on a
CC surface in a subject. The present sequence represents a hairless protein
CC interacting partner used in the composition.
XX
SQ Sequence 693 AA;

Query Match 99.4%; Score 3682; DB 8; Length 693;
Best Local Similarity 99.4%; Pred. No. 7e-297; 3; Indels 0; Gaps 0;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VTQCQSCVQAAGVGVLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
Db 1 VTQCQSCVQAAGVGVLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
Qy 61 AXHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWRC 120
Db 61 AXHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWRC 120
Qy 121 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAAGHAACSLILTQFVSSQALAE 180
Db 121 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAAGHAACSLILTQFVSSQALAE 180
Qy 181 TVMHQWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQSCNGDSNRTKDIKE 240
Db 181 TVMHQWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQSCNGDSNRTKDIKE 240
Qy 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Qy 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGLSLPLSPVTRRLSPPGALLWLQEP 360
Db 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGLSLPLSPVTRRLSPPGALLWLQEP 360
Qy 361 HGPHLFQEHWRQGPVLSVGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
Db 361 HGPHLFQEHWRQGPVLSVGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
Qy 421 FWEFGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 421 FWEFGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Qy 481 LGTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540
Db 481 LGTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540
Qy 541 GLDGEGLWSPGQSTSTVHVVFRAQDAQRIRRFQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
Db 541 GLDGEGLWSPGQSTSTVHVVFRAQDAQRIRRFQMVCPAGAGTLEPGAPGSCYLDAGLRR 600
Qy 601 RUREEHWGVCWTLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSALQCHQ 660
Db 601 RUREEHWGVCWTLQAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSALQCHQ 660
Qy 661 ASLPPDHRMLYQMDRAVFOAVKAAVNGALQEA 693
Db 661 ASLPPDHRMLYQMDRAVFOAVKAAVNGALQEA 693

RESULT 2
AAE19798
ID AAE19798 standard; protein; 1182 AA.
XX
AC AAE19798;
XX
DT 18-JUN-2002 (first entry)
XX
DE Mouse Hairless protein (Hr).
XX
KW Mouse; Hairless protein; Hr protein; Hr gene; dermatological condition;

KW hair loss; gene therapy.
XX Mus musculus.
OS US6348348-B1.
PN 19-FEB-2002.
PD 07-APR-1999; 99US-00287354.
PF 07-APR-1998; 98US-0080888P.
PR (CARN-) CARNEGIE INST WASHINGTON.
XX Thompson CC;
XX WPI; 2002-204622/26.
XX Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.
XX Example; Fig 1; 48pp; English.
XX The invention relates to human Hairless (Hr) polypeptides and nucleic acid molecules (HR) encoding such polypeptides. The invention also relates to the Hairless expression constructs which may be used in transfection assays. Sequences of the invention are used in diagnosis and treatment of dermatological conditions such as hair loss. They are also used in gene therapy. Polynucleotides of the invention can be used as probes for the detection of hair loss. The present sequence is mouse Hairless protein (Hr)
XX Sequence 1182 AA;

Query Match 99.4%; Score 3682; DB 5; Length 1182;
Best Local Similarity 99.4%; Pred. No. 1.5e-296;
Matches 689; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VTQCQSCVQAAGVGVLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 60
Db 490 VTQCQSCVQAAGVGVLTGHSQKRSRSPLEEKQLEEDSSATSEEGGGPGPEASLNKGL 549
Qy 61 AXHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWRC 120
Db 550 AXHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTEDSPGIPHCSCRCHHGLFNTHWRC 609
Qy 121 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAAGHAACSLILTQFVSSQALAE 180
Db 610 SHCSHRLCVACGRIAGAKNREKTSQEOHTDDCAQEAAGHAACSLILTQFVSSQALAE 669
Qy 181 TVMHQWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQSCNGDSNRTKDIKE 240
Db 670 TVMHQWAKFDIRGHCFQVDARVWAPGDGGQKQKTEPTPTPPQSCNGDSNRTKDIKE 729
Qy 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 730 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 789
Qy 301 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGLSLPLSPVTRRLSPPGALLWLQEP 360
Db 790 NILDSIIAQVVERKIOEKALGPGLRAGSLRKGLSLPLSPVTRRLSPPGALLWLQEP 849
Qy 361 HGPHLFQEHWRQGPVLSVGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 420
Db 850 HGPHLFQEHWRQGPVLSVGIQKTLRLSLWGMALGTLGGQVQSLTALGPPQPTNLDSTA 909
Qy 421 FWEFGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 910 FWEFGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 969
Qy 481 LGTLHPLPQLWAAVGVNSHRHGLTKNLCVEVSDLSILVHAEALQPPWYRAQKDFLS 540

970	IGITLHLEPQLWAAYGWNSHRGHGTGYNLCVEVSDLLISILVHAEQLPWYRAQKDFLS	1029
541	GLDGEGLWSPGSGTSTVWHVFRAQDAQRIIRFLQWCPAGAGTLEPGAPGSCYLDAGLR	600
1030	GLDGEGLWSPGSGTSTVWHVFRAQDAQRIIRFLQWCPAGAGTLEPGAPGSCYLDAGLR	1089
601	RLREEMGVCWMTLLQAPGEAVLPVAPGAPHVQGLVSTISVTOHFLSPETSALSACLCHQ	660
1090	RLREEMGVCWMTLLQAPGEAVLPVAPGAPHVQGLVSTISVTOHFLSPETSALSACLCHQ	1149
661	ASLPDPHRLMYAQMDRVAFQAVKAAVGAALQEA	693
1150	ASLPDPHRLMYAQMDRVAFQAVKAAVGAALQEA	1182

RESULT 3

ABU62544
ID ABU62544 standard; protein: 1182 AA.

AC ABU62544;

XX
DT 08-SEP-2003 (first entry)XX
DE
Mouse Hairless (Hr) polypeptide.

Mouse; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
 family pedigree; radiation hybrid; somatic cell hybrid.

XX
SO
Mus sp.

XX
PN
US2003027300-A1.XX
PD 06-FEB-2003.

XX
PF 21-DEC-2001; 2001US-00024368.

XX
PR 07-APR-1998; 98US-0080888P.

PK	07-APR-1998;	98US-0080888F.
PR	07-APR-1999;	99US-00287354.

XX PA (CARN-) CARNEGIE INST WASHINGTON.

XX
PI
Thompson CC;XX
DR WPI: 2003-492034/58.

xx PT New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.

XX Example; Fig 1; 34pp; English. PS

The invention relates to a human hairless (HR) polynucleotide encoding the hairless (Hr) polypeptide. The invention also relates to an expression system comprising an expression construct which produces a polypeptide with hairless transcription factor activity, a reporter construct comprising a transcription regulatory region responsive to hairless transcription activity to regulate transcription of the reporter gene which is mediated by the transcription regulatory region and methods of screening for chemical agents which modulate hairless-mediated transcription, binding between hairless and thyroid hormone receptor or hairless activity. The polynucleotide is useful as a probe or primer to quantitate cognate RNA and DNA within cells, which can be subsequently used to correlate hair growth or loss with hairless expression or hairless-regulated transcription. The amino acid sequence of hairless antigen can be used for preparing specific binding molecules (e.g. polyclonal or monoclonal antibodies) for monitoring protein expression, for affinity purification and for functional studies. The human hairless polynucleotide, polypeptide or specific binding molecule may be used to identify and detect this genetic marker in family pedigrees, radiation hybrids or human-rodent somatic cell hybrids, and as an affinity tag to identify, isolate and detect interacting proteins that bind the hairless gene or protein. This sequence represents a mouse hairless clypeptide of

PR	07-APR-1998;	98US-0080888P.
XX	(CARN-) CARNEGIE INST WASHINGTON.	
PA	Thompson CC;	
XX	WPI; 2002-204622/26.	
XX	Novel expression construct, useful in the diagnosis and treatment of dermatological conditions, such as hair loss, contains a Hairless gene sequence.	
PT	Example; Fig 1; 48pp; English.	
XX	The invention relates to human Hairless (Hr) polypeptides and nucleic acid molecules (HR) encoding such polypeptides. The invention also relates to the Hairless expression constructs which may be used in transcription assays. Sequences of the invention are used in diagnosis and treatment of dermatological conditions such as hair loss. They are also used in gene therapy. Polynucleotides of the invention can be used as probes for the detection of hair loss. The present sequence is human Hairless protein	
XX	SQ	Sequence 1189 AA;
Query Match	83.6%;	Score 3095.5; DB 5; Length 1189;
Best Local Similarity	83.2%;	Pred. No. 9.5e-248;
Matches	579;	Conservative 40; Mismatches 74; Indels 3; Gaps 2;
Qy	1	VTCQSCVQAAGEVGLTGHSHQSKRRSPL-EEKQLEEDSSSATSEBGGGPGPEASLNKG 59
Db	494	LAQCSCQAAGGEGGHACHSQVRRSPGLGELQOEEDTATNSSSEBGGGPGDRLSTG 553
Qy	60	LAKHLLSGLDRLCLLRKEREALAWAQREGQGPANTEDESPGPHCCSRCHHGLFNTHWR 119
Db	554	LAKHLLSGLDRLCLLREREALAWAQREGQGPANTEDESPGPRCCSRCHHGLFNTHWR 613
Qy	120	CSCSHRLCVACGRIAGAGKNREKTSQEOTDDCAQEAAGHAACSLILTQFVSSQALAE 179
Db	614	CPRCSHRLCVACGRVAGTGEAREKAGFQGSAECTQEAAGHAACSLMLTQFVSSQALAE 673
Qy	180	STVMHQAWKFDIRGHCFQVDARVWAPGGGQOQKEPTKPTPTPSCNGDSNRTKD 239
Db	674	STAMHQVWVKFDIRGHCFQVDARVWAPGDAGQOQKESTQKPTPTPSCNGDTHRTKSIK 733
Qy	240	EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRTHMAFAPVTPALPSDDRI 299
Db	734	EETPDSAETPAEDRAGRFLPCPSLCELLASTAVKLCIGHERTHMAFAPVTPALPSDDRI 793
Qy	300	TNILDSTIAQVVERKIQEALGPGRAGSLRKLPLSPVTRTSLPPGALLWLOEPR - 358
Db	794	TNILDSTIAQVVERKIQEALGPGRAGSLRKLPLSPVTRTSLPPGALLWLOEPR - 853
Qy	359	-PKXGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTGLGQVQSLTALGPPQPTNLD 417
Db	854	CPRRGFHLFQEHWRQGPVLVSGIQRTLQGNLWGTALGALGQVQALSPLGPPQSSLG 913
Qy	418	STAFWEGFSPHETPRPKLDEGSVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLAS 477
Db	914	STTFWEGFSPHETPRPKSDEGSVLLHRLALGDEDTSRVENLAASLPPEYCALHGLKNLAS 973
Qy	478	YLPGLGLTLPLEPQLWAAVGVNSHRHGLGTKNLCVEVSDLSITLVHAEALQPLPWYRAOKD 537
Db	974	YLPGLGLTLPLEPQLWAAVGVNSHRHGLGTKNLCVEVADVLVLLVHADTPLPAWHRAOKD 1033
Qy	538	FLSLGDLGELWSPGSGTSTVWHVFRQAQORIRRFLOWVCPAGAGTLEPGAGSCYLDAG 597
Db	1034	FLSLGDLGELWSPGSGVSTVWHVFRQAQORIRRFLOWVCPAGAGALEPGAGSCYLDAG 1093
Qy	598	LRRLREEWGVCWTLTLOAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 657
Db	1094	LRRLREEWGVCWTLTLOAPGEAVLPAGAPHQVQGLVSTVSTQHFLSPETSALSQALC 1153

QY 658 HQGASLPDHRMLYQMDRAVFOAVKAAVCAALQAEK 693
 DB 1154 HQGPSLPDCHLLYQMDWAQVFOAVKAVGTLLQAEK 1189

RESULT 7
 ABU62542
 ID ABU62542 standard; protein; 1189 AA.
 AC ABU62542;
 DT 08-SEP-2003 (first entry)
 XX Human Hairless (Hr) polypeptide #4.
 XX Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
 KW family pedigree; radiation hybrid; somatic cell hybrid.
 XX Homo sapiens.
 OS US2003027300-A1.
 PN 06-FEB-2003.
 XX 21-DEC-2001; 2001US-00024368.
 PF 07-APR-1998; 98US-0080888P.
 PR 07-APR-1999; 99US-00287354.
 PA (CARN-) CARNEGIE INST WASHINGTON.
 XX Thompson CC;
 PI WPI; 2003-492034/58.
 DR New Hairless polypeptide and polynucleotide, useful for identifying and
 PT detecting this genetic marker in family pedigrees or human-rodent somatic
 PT cell hybrids, or detecting interacting proteins that bind hairless gene
 PT or protein.
 XX Example; Fig 1; 34pp; English.
 XX The invention relates to a human hairless (HR) polynucleotide encoding
 CC the hairless (HR) polypeptide. The invention also relates to an
 CC expression system comprising an expression construct which produces a
 CC polypeptide with hairless transcription factor activity, a reporter
 CC construct comprising a transcription regulatory region responsive to
 CC hairless transcription activity to regulate transcription of the reporter
 CC gene which is mediated by the transcription regulatory region and methods
 CC of screening for chemical agents which modulate hairless-mediated
 CC transcription, binding between hairless and thyroid hormone receptor or
 CC hairless activity. The polynucleotide is useful as a probe or primer to
 CC quantitate cognate RNA and DNA within cells, which can be subsequently
 CC used to correlate hair growth or loss with hairless expression or
 CC hairless-regulated transcription. The amino acid sequence of hairless
 CC antigen can be used for preparing specific binding molecules (e.g.
 CC polyclonal or monoclonal antibodies) for monitoring protein expression,
 CC for affinity purification and for functional studies. The human hairless
 CC polynucleotide, polypeptide or specific binding molecule may be used to
 CC identify and detect this genetic marker in family pedigrees, radiation
 CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
 CC identify, isolate and detect interacting proteins that bind the hairless
 CC gene or protein. This sequence represents a human hairless polypeptide of
 CC the invention
 XX Sequence 1189 AA;

Query Match 83.6%; Score 3095.5; DB 6; Length 1189;
 Best Local Similarity 83.2%; Pred. No. 9.5e-248;
 Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
 QY 1 VTCQSCVQAGEVGLTGHISQKRRRLP-EEKQLEEDSSATSEGGGPGPEASLNKG 59

DB 494 LAQCSCAQAGEGGGHACHSQVRRSPGLGELQEEEDTATNSSSEEGPGSPDSRLSTG 553
 QY 60 LAKHLLSGLDRLCRLLRKRERALAWAQREGGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 119
 DB 554 LAKHLLSGLDRLCRLLRKRERALAWAQREGGQGPAMTEDSPGPHCCSRCHHGLFNTHWR 613
 QY 120 CSHCSHRLCVACGRIAGCKNREKTSQEOHTDDCAQEAAGHACSLILTFQVSSQALAE 179
 DB 614 CPRCSHRLCVACGRVAGTGRAREKAGFQESAECEQEAAGHACSLILTFQVSSQALAE 673
 QY 180 STVMHQAQAFDIRGHCFQVDARVWAPGDGGQOKEPTEKTPPTPSCNGDSNRKDKIK 239
 DB 674 STAMHQQVWVWFDIRGHCFQVDARVWAPGDGGQOKEPTEKTPPTPSCNGDTHRTSIK 733
 QY 240 EETPDSTESPADGAGRSPLPCPSICELLASTAVKLCGLGHDRIHMAFAPVTPALPDDRI 299
 DB 734 EETPDSAETPAEDRAGRPLPCPSICELLASTAVKLCGLGHDRIHMAFAPVTPALPDDRI 793
 QY 300 TNILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRTLSPPCALLWLQEP 358
 DB 794 TNILDSIIAQVVERKIQEALGPGLRAGSLRGLSLPLSPVTRTLSPPCALLWLQEP 853
 QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTLGQGVQSLTALGPPPTNLD 417
 DB 854 CPERGHLFQEHWRQGPVLVSGIQKTLRLSLMGMEALGTLGQGVQSLTALGPPPTNLD 913
 QY 418 STAFWEGFHPETRPKLDGSGVLLHRTLDGDASRVQNLVSSLPPLPEYCAHOGKLNLAS 477
 DB 914 STTFWEGFSWPELRPKSDGSGVLLHRTLDGDASRVQNLVSSLPPLPEYCAHOGKLNLAS 973
 QY 478 YLPLGLTLHPLPOLWAAVGVNSHRHGLTKNLCEVSDLSILVHAEALQPPWYRAQKD 537
 DB 974 YLPPGLALRPLEPQLWAAVGVNSHRHGLTKNLCEVSDLSILVHAEALQPPWYRAQKD 1033
 QY 538 FLGSLDGEGLWSPGSGTSTVWHVFRQAQRIIRRFQWVCPAGAGTLEPGAGSCYLDAG 597
 DB 1034 FLGSLDGEGLWSPGSGTSTVWHVFRQAQRIIRRFQWVCPAGAGTLEPGAGSCYLDAG 1093
 QY 598 LRRRLREWGVCWTLLOAPGEAVLPAGAPHQVGLVSTISVTQHFLSPETSALSALQC 657
 DB 1094 LRRRLREWGVCWTLLOAPGEAVLPAGAPHQVGLVSTISVTQHFLSPETSALSALQC 1153

658 HQGASLPDHRMLYQMDRAVFOAVKAAVCAALQAEK 693
 1154 HQGPSLPDCHLLYQMDWAQVFOAVKAVGTLLQAEK 1189

RESULT 8
 ID AAY15218 standard; protein; 1189 AA.
 AC AAY15218;
 XX AAY15218;
 DT 09-NOV-1999 (first entry)
 XX Human Hairless mutant amino acid sequence, Thr(1022)Ala.
 XX alopecia; congenital alopecia; congenital atrichia;
 KW androgenetic alopecia; alopecia areata; alopecia universalis; mutant;
 KW hair follicle.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1022
 FT /label= T1022A
 FT /note= "wild-type Thr substituted by Ala"
 XX WO9938965-A1.
 XX 05-AUG-1999.
 XX 29-JAN-1999; 99WO-US002128.

Qy	60	LAKHLLSGLDRLCRLLRKREALAWAQROGGPAMTDESPGI	PHCCSRCHHGLFNTHWR	119
Db	554	LAKHLLSGLDRLCRLLRREAREALAWAQRESQGPVTE	SPGIPRCCSRCHHGLFNTHWR	613
Qy	120	CSHSHRLCVACGRICAGAKNREKTKTQEOHTDDCAOEAGHAACSL	LLTLQFVSSQALAE	179
Db	614	CPRCSHRLCVACGRVAGTGRAREKAGFQEOSAEECTQEA	HAACSLMLTQFVSSQALAE	673
Qy	180	STVMHQAWKFDIRGHCFQVDARVWAPGGGQOKEPTKTP	PPQSPSCNGDSNRTKDIK	239
Db	674	STAMHQVWVKFDIRGHCFQCADARVWAPGDAQOKESTQ	KTPPTPPQSPSCNGDTHRTKSIK	733
Qy	240	EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLC	LGHDRIHMAFAPVTPALPSDDRI	299
Db	734	EETPDSAETPAEDRAGRGPLPCPSLCELLASTAVKLC	LGHERIHMAFAPVTPALPSDDRI	793
Qy	300	TNILDSTIIAQVVERKIOEKALGPGLRAGSLRGKLS	PLSPVTRTLPSPGALLWLOEPR	358
Db	794	TNILDSTIIAQVVERKIOEKALGPGLRAGSLRGKLS	PLSPVTRTLPSPGALLWLOEPPQ	853
Qy	359	- PKHGPHLFOBHWRGQGPVLVSGIQTKTLRLSLWQNEAL	GTGGVQSLTALGPPQPTNLD	417
Db	854	CPRRGFFLFOBHWRGQGPVLVSGIQTKTLQGNLWGT	EALGALGGVQVQALSLGPPQPSL	913
Qy	418	STARFEGFSPEPTEPKLDEGSLVLLHRTLDGDKDASR	VQNLVSSLPPLPEYCAHOGKUNLAS	477
Db	914	STTFWEGFSPELPKPKDEGSLVLLHRLALGDEDT	SRVENLAASLPPEYCALHOGKUNLAS	973
Qy	478	YLPGLGLTLHPLEPOLWAAAYGNSHRHGLTKNLK	CEVSDLTISILVHAEAOQLPPMYRAQKD	537
Db	974	YLPGLGLTLHPLEPOLWAAAYGVS	PHRGHLGTKNLCEVADLVSLVHADTPLPAMHRAQKD	1033
Qy	538	FLSGLDGEGLSWPGSQSTSTVWHVFRADOAQRIR	RFLOWCPAGAGTLEPGAPGSCYLDAG	597
Db	1034	FLSGLDGEGLSWPGSQSTSTVWHVFRADOAQRIR	RFLOWCPAGAGALEPGAPGSCYLDAG	1093
Qy	598	LRRRLREEWGVS	CWTLTLOAQCEAVLVPAGAPHOVQGLVSTISVTOHFLSPETSALSQLC	657
Db	1094	LRRRLREEWGVS	CWTLTLOAQCEAVLVPAGAPHOVQGLVSTISVTOHFLSPETSALSQLC	1153
Qy	658	HQGLASLPDPDRMLYQAOMDRAVFQAVKAAVGALQ	BEAK	693
Db	1154	HQGLASLPDPDCHLLYQAOMDWAVFQAVKAVGTLQ	BEAK	1189

RESULT 10	
ABG72775	standard; protein; 1189 AA.
ID	ABG72775
XX	
AC	ABG72775;
XX	
DT	25-FEB-2003 (first entry)
XX	
DE	Human hairless protein.
XX	
KW	Human; catalytic DNA; catalytic RNA; hairless protein; hair loss;
KX	atrachia; hair growth; hirsutism; catalytic nucleic acid; ribozyme;
KW	DNAzyme; self-catalytic; hammerhead ribozyme; deoxy-ribozyme;
KW	catalytic core; cleavage site; pharmaceutical; hair production;
KW	hair follicle; anagen phase; catagen phase; hair removal product;
KW	depilatory.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 446 /note= "Encoded by TGG"
FT	
FT	Misc-difference 584 /note= "Encoded by GGC"
FT	
XX	
PX	WO200283891-A2.
XX	
PD	24-OCT-2002.

Db 854 CPERGFHLFOEHWROQPVLSGIQRTLOGNLWGTALGALGQVQALSPGPPQSSLG 913
QY 418 STAFWEGFSGHPETPKLDEGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLS 477
Db 914 STTFWEGFSGWPELPRKSDSGSVLLHRLGDEDTSRVENLAASLPLPEYCALHKGKLNLS 973
QY 478 YLPLGLTLHPLEPOLWAAVGVNSHRHGLGTGKNCVEVSDLSILVHAELQPLPPYRAQKD 537
Db 974 YLPPGLALRPLEPOLWAAVGVNSHRHGLGTGKNCVEVADLVSLVHADTPLPAWHRAQKD 1033
QY 538 FLUGLDEGLWSPGSGTSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAG 597
Db 1034 FLUGLDEGLWSPGSGVSTVWHVFRADQRIIRRFLOWMCPAGAGALEPGAGSCYLDAG 1093
QY 598 LRRRLREMGVSCWTLLOAPGEAVLPAGAPHOVQGLVSTVSTVQHFLSPETSALSQAOLC 657
Db 1094 LRRRLREMGVSCWTLLOAPGEAVLPAGAPHOVQGLVSTVSTVQHFLSPETSALSQAOLC 1153
QY 658 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 693
Db 1154 HQGPSLPDPCHLLYQAQMDWAVFAVKAAGVAVGTLQAEK 1189

RESULT 11

AAE19794
ID AAE19794 standard; protein; 984 AA.

AC AAE19794;

DT 18-JUN-2002 (first entry)

XX Human Hairless protein (Hr) #1.

XX Human; Hairless protein; Hr protein; Hr gene; dermatological condition;
KW hair loss; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 284
/note= "Encoded by CTT of the sequence shown as SEQ ID
NO:1 (AAD31072) in Column 35-38 of the specification"

FT Misc-difference 978

FT /note= "Encoded by CGG of the sequence shown as SEQ ID
NO:1 (AAD31077) in Column 25-32 of the specification"

FT Misc-difference 982

FT /note= "Encoded by CAG of the sequence shown as SEQ ID
NO:1 (AAD31077) in Column 25-32 of the specification"

XX US6348348-B1.

XX 19-FEB-2002.

XX 07-APR-1999; 99US-00287354.

XX 07-APR-1998; 98US-0080888P.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX Thompson CC;

XX WPI; 2002-204622/26.

XX N-PSDB; AAD31072, AAD31077.

XX Novel expression construct, useful in the diagnosis and treatment of
XX dermatological conditions, such as hair loss, contains a Hairless gene
XX sequence.

XX Claim 1b; Fig 1; 48pp; English.

XX The invention relates to human Hairless (Hr) polypeptides and nucleic
XX acid molecules (HR) encoding such polypeptides. The invention also

CC relates to the Hairless expression constructs which may be used in
CC transcription assays. Sequences of the invention are used in diagnosis
CC and treatment of dermatological conditions such as hair loss. They are
CC also used in gene therapy. Polynucleotides of the invention can be used
CC as probes for the detection of hair loss. The present sequence is human
CC Hairless protein (Hr)

XX Sequence 984 AA;

Query Match 83.4%; Score 3088.5; DB 5; Length 984;

Best Local Similarity 83.0%; Pred. No. 2.8e-247;

Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHSGKSRSPLE-ERKQLEEDSSATSBEGGGGPGPEASLNKG 59

Db 289 LAQCSCQAAGEGGCHACHSQVRRSPGLGELQEDDTATNSSSEGGSGDPDSRLSTG 348

QY 60 LAKHLLSGLDRLCLRLRERREALAWAQREGGQPMANEDSPGPHCCSCHHGLFNTHWR 119

Db 349 LAKHLLSGLDRLCLRLRERREALAWAQREGGQPAVTGDSPIRCCSCHHGLFNTHWR 408

QY 120 CSHCSHRLCVACRTAGAKNREKTSQEOHTDDCAQEAHAACSLILTFVSSQAALAE 179

Db 409 CPRCSHRLCVACRGVAGTGRAREKAGFQESABECTQEAGHAACSLMLTFQVSSQAALAE 468

QY 180 STVMHQAWAKFDIRGHCFQVDARVWAPGDGGQOKEPTKPTPTPOPCNGSDNRKTDIK 239

Db 469 STAMHQVWVKFDIRGHCPQADARVWAPGDAGQOKESTOKTPTPTPOPCNGDTHRTSIK 528

QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPDDRI 299

Db 529 EETPDSAETPAEDRAGRGPLPCPSLCELLASTAVKLCIGHDRHMAFAPVTPALPDDRI 588

QY 300 THILDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEP- 358

Db 589 THILDSIIAQVVERKIQEALGFLRAGSLRKLGLSLPLSPVTRLSPPGALLWLQEPQP 648

QY 359 -PKHGFHLFOEHWROQPVLSGIQKTLRLSLWMEALGTGGOVQSLTALGPPQPTNLD 417

Db 649 CPERGFHLFOEHWROQPVLSGIQKTLQGNLWGTALGALGQVQALSPGPPQSSLG 708

QY 418 STAFWEGFSGHPETPKLDEGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHOGKLNLS 477

Db 709 STTFWEGFSGWPELPRKSDSGSVLLHRLGDEDTSRVENLAASLPLPEYCALHKGKLNLS 768

QY 478 YLPLGLTLHPLEPOLWAAVGVNSHRHGLGTGKNCVEVSDLSILVHAELQPLPPYRAQKD 537

Db 769 YLPPGLALRPLEPOLWAAVGVNSHRHGLGTGKNCVEVADLVSLVHADTPLPAWHRAQKD 828

QY 538 FLUGLDEGLWSPGSGTSTVWHVFRADQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAG 597

Db 829 FLUGLDEGLWSPGSGVSTVWHVFRADQRIIRRFLOWMCPAGAGALEPGAGSCYLDAG 888

QY 598 LRRRLREMGVSCWTLLOAPGEAVLPAGAPHOVQGLVSTVSTVQHFLSPETSALSQAOLC 657

Db 889 LRRRLREMGVSCWTLLOAPGEAVLPAGAPHOVQGLVSTVSTVQHFLSPETSALSQAOLC 948

QY 658 HQGASLPDPHRLMYAQMDRAVFAVKAAGALQAEK 693

Db 949 HQGPSLPDPCHLLYQAQMDWAVFAVKAAGVAVGTLQAEK 984

RESULT 12

ABU62545

ID ABU62545 standard; protein; 984 AA.

XX AC ABU62545;

XX DT 08-SEP-2003 (first entry)

XX DE Human Hairless (Hr) polypeptide #2.

XX KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;


```
CC putative single zinc finger transcription factor protein (Hairless)
XX
SQ Sequence 1189 AA;

Query Match      83.1%; Score 3077.5; DB 5; Length 1189;
Best Local Similarity 82.9%; Pred. No. 3e-246;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGVGVLTHGSHQSRSPLE-EEKLEEDSSATSEGGGGPGPEASLNKG 59
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 LAQCSCQAAGGEGGHACHSQQVRRSPGLGELQOEEDTATNSSSEGGPGDRLSTG 553
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 60 LAKHLLSGLGDLRLCLLRKEREALAWAQREGQGPAMTDSFGPHCCSRCHHGLFNTHWR 119
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 LAKHLLSGLGDLRLCLLRKEREALAWAQREGQGPAMTDSFGPHCCSRCHHGLFNTHWR 613
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTFVSSQALAE 179
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 CSHCSHRLCVACGRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTFVSSQALAE 673
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 180 STVMHQAWAKFDIRGHCFQVDARVWAPGDGQOQKEPTKTPTTPOPCNGDSNRKTDIK 239
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 STAHQVWVAFDIRGHCFQVDARVWAPGDGQOQKEPTKTPTTPOPCNGDTHRTKSIK 733
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 EETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 300 TNLDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEP- 358
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 TNLDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEP- 853
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 359 -PKHGPHLFOEHWKQGPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 417
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 CPRGPHLFOEHWKQGPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 913
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 418 STAFWEGFHPETPKLDEGSLVLLHRTLDGDKDASRVQNLVSLPLPEYCAHQKLNLAS 477
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 STTFWEGFHPETPKLDEGSLVLLHRTLDGDKDASRVQNLVSLPLPEYCAHQKLNLAS 973
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 478 YLPLGLTLHLEPOLWAAVGNHSHGLGTHKNLCVEVSDILSVLVAEALPWPYRAQKD 537
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 YLPPGLALPPLWAAVGNHSHGLGTHKNLCVEVSDILSVLVAEALPWPYRAQKD 1033
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 538 FLGSLDGEGLWSPGSGSTVWHVFRQDAQRIIRFLQWCPAGAGTLEPGAGSCVLDAG 597
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 FLGSLDGEGLWSPGSGSTVWHVFRQDAQRIIRFLQWCPAGAGTLEPGAGSCVLDAG 1093
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 598 LRRRLREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 657
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 LRRRLREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 1153
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 658 HQGASLPDPHMLYIAQMDRAVFAVKAVALQEAQ 693
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 HQGASLPDPHMLYIAQMDRAVFAVKAVALQEAQ 1189
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
ABU62541
ID ABU62541 standard; protein; 1189 AA.
XX
AC ABU62541;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human Hairless (Hr) polypeptide #3.
XX
KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
XX family pedigree; radiation hybrid; somatic cell hybrid.
XX
OS Homo sapiens.
XX
PN US2003027300-A1.
```

```
XX
PD
XX
PF 21-DEC-2001; 2001US-00024368.
XX
PR 07-APR-1998; 98US-0080888P.
XX 07-APR-1999; 99US-00287354.
XX (CARN-) CARNEGIE INST WASHINGTON.
PA
PI Thompson CC;
XX
XX WPI; 2003-492034/58.
XX
XX New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
PT cell hybrids, or detecting interacting proteins that bind hairless gene
PT or protein.
XX
PS Example; Fig 1; 34pp; English.
XX
XX The invention relates to a human hairless (HR) polynucleotide encoding
CC the hairless (Hr) polypeptide. The invention also relates to an
CC expression system comprising an expression construct which produces a
CC polypeptide with hairless transcription factor activity, a reporter
CC construct comprising a transcription regulatory region responsive to
CC hairless transcription activity to regulate transcription of the reporter
CC gene which is mediated by the transcription regulatory region and methods
CC of screening for chemical agents which modulate hairless-mediated
CC transcription, binding between hairless and thyroid hormone receptor or
CC hairless activity. The polynucleotide is useful as a probe or primer to
CC quantitate cognate RNA and DNA within cells, which can be subsequently
CC used to correlate hair growth or loss with hairless expression or
CC hairless-regulated transcription. The amino acid sequence of hairless
CC antigen can be used for preparing specific binding molecules (e.g.
CC polyclonal or monoclonal antibodies) for monitoring protein expression,
CC for affinity purification and for functional studies. The human hairless
CC polynucleotide, polypeptide or specific binding molecule may be used to
CC identify and detect this genetic marker in family pedigrees, radiation
CC hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
CC identify, isolate and detect interacting proteins that bind the hairless
CC gene or protein. This sequence represents a human hairless polypeptide of
CC the invention
XX
SQ Sequence 1189 AA;

Query Match      83.1%; Score 3077.5; DB 6; Length 1189;
Best Local Similarity 82.9%; Pred. No. 3e-246;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGVGVLTHGSHQSRSPLE-EEKLEEDSSATSEGGGGPGPEASLNKG 59
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 LAQCSCQAAGGEGGHACHSQQVRRSPGLGELQOEEDTATNSSSEGGPGDRLSTG 553
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 60 LAKHLLSGLGDLRLCLLRKEREALAWAQREGQGPAMTDSFGPHCCSRCHHGLFNTHWR 119
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 554 LAKHLLSGLGDLRLCLLRKEREALAWAQREGQGPAMTDSFGPHCCSRCHHGLFNTHWR 613
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 120 CSHCSHRLCVACGRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTFVSSQALAE 179
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 614 CSHCSHRLCVACGRVAGTGRAREKAGFQESAEECTQEAAGAACSLMLTFVSSQALAE 673
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 180 STVMHQAWAKFDIRGHCFQVDARVWAPGDGQOQKEPTKTPTTPOPCNGDSNRKTDIK 239
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 674 STAHQVWVAFDIRGHCFQVDARVWAPGDGQOQKEPTKTPTTPOPCNGDTHRTKSIK 733
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 734 EETPDSAETPAEDRAGRGLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 300 TNLDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEP- 358
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 TNLDSIIAQVVERKIQEAKGLPGLRAGSLRGLSLPLSPVTRLSPPGALLWLQEP- 853
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 359 -PKHGPHLFOEHWKQGPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 417
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 CPRGPHLFOEHWKQGPVLVSGIQKTLRLSLMGMEALGTGQVQSLTALGPPQPTNLD 913
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 418 STAFWEGFHPETPKLDEGSLVLLHRTLDGDKDASRVQNLVSLPLPEYCAHQKLNLAS 477
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 STTFWEGFHPETPKLDEGSLVLLHRTLDGDKDASRVQNLVSLPLPEYCAHQKLNLAS 973
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 478 YLPLGLTLHLEPOLWAAVGNHSHGLGTHKNLCVEVSDILSVLVAEALPWPYRAQKD 537
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 YLPPGLALPPLWAAVGNHSHGLGTHKNLCVEVSDILSVLVAEALPWPYRAQKD 1033
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 538 FLGSLDGEGLWSPGSGSTVWHVFRQDAQRIIRFLQWCPAGAGTLEPGAGSCVLDAG 597
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 FLGSLDGEGLWSPGSGSTVWHVFRQDAQRIIRFLQWCPAGAGTLEPGAGSCVLDAG 1093
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 598 LRRRLREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 657
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 LRRRLREEMGVSCWTLQAPGEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSQALC 1153
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 658 HQGASLPDPHMLYIAQMDRAVFAVKAVALQEAQ 693
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 HQGASLPDPHMLYIAQMDRAVFAVKAVALQEAQ 1189
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWGMALGTLGGVQVSLTALGPPQPTNLD 417
Db 854 CRRGRFHLFQEHWRQGPVLVSGIQLTQGNLWGTALGALGGVQVQALSPLAPPQSSLG 913
QY 418 STAFWEGFHPETRPKLDGSGVLLHRTLDGDKASRNVNLSLPLPEYCAHOGKLNLS 477
Db 914 STTFWEGFHPETRPKLDGSGVLLHRTLDGDKASRNVNLSLPLPEYCAHOGKLNLS 973
QY 478 YLPLGLTLHPLEPQLWAAVGNHGHGTLGKNCVSVSLILVHAEOALPWPYRAQKD 537
Db 974 YLPPGLALRLEPQLWAAVGNHGHGTLGKNCVSVADLVSLVHARTPLPAWHAQKD 1033
QY 538 FLSGLDGGLWSPGSGTSTVHVFRQAQDQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAG 597
Db 1034 FLSGLDGGLWSPGSGTSTVHVFRQAQDQRIIRFLQWVCPAGAGTLEPGAPGSCYLDAG 1093
QY 598 LRRRLREWGVSWTLLQAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALQC 657
Db 1094 LRRRLREWGVSWTLLQAPGEAVLPAGAPHQVQGLVSTVTOHFLSPETSALSALQC 1153
QY 658 HQGASLPDPHRLMYAQMDFVFAVKAAGVCAQLEAK 693
Db 1154 HQGASLPDPCHLLYQAQMDWAVFAVKAAGVCAQLEAK 1189
RESULT 15
ABU62540
ID ABU62540 standard; protein; 984 AA.
AC ABU62540;
XX
DT 08-SEP-2003 (first entry)
XX
DE Human Hairless (Hr) polypeptide #1.
KW Human; hairless; Hr; thyroid hormone receptor; hair growth; hair loss;
XX family pedigree; radiation hybrid; somatic cell hybrid.
OS Homo sapiens.
FH
FT Key Location/Qualifiers
FT Misc-difference 50 /note= "Encoded by ATC"
FT Misc-difference 140 /note= "Encoded by CAG"
FT Misc-difference 384 /note= "Encoded by CTG"
FT Misc-difference 446 /note= "Encoded by GAG"
FT Misc-difference 453 /note= "Residue printed as Her"
FT Misc-difference 467 /note= "Encoded by GAG"
FT Misc-difference 529..530 /note= "Encoded by GAGGAG"
FT Misc-difference 603 /note= "Encoded by AAC"
FT Misc-difference 606 /note= "Encoded by GAG"
FT Misc-difference 645 /note= "Encoded by GAG"
FT Misc-difference 649 /note= "Residue printed as Gys"
FT Misc-difference 659 /note= "Encoded by GAG"
FT Misc-difference 731 /note= "Residue printed as Leo"
FT Misc-difference 747 /note= "Encoded by AAG"
XX
US2003027300-A1.
XX

PD 06-FEB-2003.
XX
PF 21-DEC-2001; 2001US-00024368.
XX
PR 07-APR-1998; 98US-0080888P.
PR 07-APR-1999; 99US-00287354.
XX
PA (CARN-) CARNEGIE INST WASHINGTON.
XX
PI Thompson CC;
XX
WI: 2003-492034/58.
DR N-PSDB; AC266351.
XX
New Hairless polypeptide and polynucleotide, useful for identifying and
PT detecting this genetic marker in family pedigrees or human-rodent somatic
cell hybrids, or detecting interacting proteins that bind hairless gene
or protein.
XX
PS Claim 1; Page 14-17; 34pp; English.
XX
CC The invention relates to a human hairless (HR) polynucleotide encoding
the hairless (Hr) polypeptide. The invention also relates to an
expression system comprising an expression construct which produces a
polypeptide with hairless transcription factor activity, a reporter
construct comprising a transcription regulatory region responsive to
hairless transcription activity to regulate transcription of the reporter
gene which is mediated by the transcription regulatory region and methods
of screening for chemical agents which modulate hairless-mediated
transcription, binding between hairless and thyroid hormone receptor or
hairless activity. The polynucleotide is useful as a probe or primer to
quantitate cognate RNA and DNA within cells, which can be subsequently
used to correlate hair growth or loss with hairless expression or
hairless-regulated transcription. The amino acid sequence of hairless
antigen can be used for preparing specific binding molecules (e.g.
polyclonal or monoclonal antibodies) for monitoring protein expression,
for affinity purification and for functional studies. The human hairless
polynucleotide, polypeptide or specific binding molecule may be used to
identify and detect this genetic marker in family pedigrees, radiation
hybrids or human-rodent somatic cell hybrids, and as an affinity tag to
identify, isolate and detect interacting proteins that bind the hairless
gene or protein. This sequence represents a human hairless polypeptide of
the invention
XX
SQ Sequence 984 AA;
Query Match 82.8%; Score 3067.5; DB 6; Length 984;
Best Local Similarity 82.0%; Pred. No. 1.5e-245;
Matches 571; Conservative 47; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGVGVLTGHSQKRRSPL-EEKQLEEDSSATSEGGGPGPEASLNKG 59
Db 289 LAQCSCQAAGGEGGHACHSQVRRSPLGELQEBEDTATNSSEEGPGSPDRLSTG 348
QY 60 LAKHLGLGLDRLCLRLREREAQVAGQGPAMTEDSPGPHCCSRCHGLFNTHWR 119
Db 349 LAKHLGLGLDRLCLRLREREAQVAGQGPAMTEDSPGPHCCSRCHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRIAGAGKNREKTSQEQTDDCAQEAAGAACSLTLTQVSSQALAE 179
Db 409 CPRCSHRLCVACGRIAGAGKNREKTSQEQTDDCAQEAAGAACSLTLTQVSSQALAE 468
QY 180 STVHQWAKFDIRGHCFQVDARVAPAGGGQKQEPTEKTPPTQPSGNGDSNRTDIK 239
Db 469 STAHQVWVKFDIRGHCFQVDARVAPAGGGQKQEPTEKTPPTQPSGNGDSNRTDIK 528
QY 240 EETPDSTESPAEDGAGSPLPCSLCELLASTAVKLCIGHDRHMAFPVTPALPSDDRI 299
Db 529 QQTTPDSTESPAEDGAGSPLPCSLCELLASTAVKLCIGHDRHMAFPVTPALPSDDRI 588
QY 300 TNLDSITIAQVVERKIQKALGFLGRAGSLRKLGLSLPSVTRLSPPGALLWLOPQP- 358
Db 589 TNLDSITIAQVVERKIQKALGFLGRAGSLRKLGLSLPSVTRLSPPGALLWLOPQP 648

This Page Blank (Page 2)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:00:26 ; Search time 48 seconds
(without alignments)
1193.630 Million cell updates/sec

Title: US-10-712-629B-18
Perfect score: 3704
Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAKAAGALQEAQ 693

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3682	99.4	1182	2	US-09-287-354-6
2	3510	94.8	1207	2	US-09-287-354-5
3	3095.5	83.6	1189	2	US-09-287-354-4
4	3095.5	83.6	1189	2	US-09-949-016-6931
5	3088.5	83.4	984	2	US-09-287-354-2
6	3077.5	83.1	1189	2	US-09-287-354-3
7	2333.5	63.0	943	2	US-09-949-016-7891
8	182	4.9	177	2	US-09-640-211A-1138
9	152.5	4.1	2442	2	US-09-514-247A-10
10	152.5	4.1	2442	2	US-09-538-092-1370
11	146.5	4.0	2414	2	US-08-227-536-2
12	146.5	4.0	2414	2	US-09-538-092-1289
13	146.5	4.0	2414	4	PCT-US95-04682-2
14	145.5	3.9	2441	1	US-08-194-468-2
15	145.5	3.9	2441	2	US-08-961-739-2
16	145.5	3.9	2441	2	US-09-514-247A-8
17	145.5	3.9	2441	2	US-09-686-316-2
18	127.5	3.4	1479	2	US-08-840-062-2
19	124.5	3.4	427	2	US-09-502-540-10191
20	121	3.3	732	1	US-08-317-522A-5
21	121	3.3	778	1	US-08-439-818A-5
22	121	3.3	778	1	US-08-751-965-5
23	121	3.3	778	1	US-08-738-975-5
24	121	3.3	778	1	US-08-728-626-5
25	121	3.3	778	1	US-08-808-599A-5
26	121	3.3	987	2	US-09-252-991A-27483
27	120	3.2	861	2	US-09-252-991A-18375

28	118	3.2	656	2	US-09-302-540-12404	Sequence 12404, A
29	117.5	3.2	662	2	US-10-104-047-3398	Sequence 3398, Ap
30	116.5	3.1	2157	2	US-09-466-778-2	Sequence 2, Appli
31	115.5	3.1	574	2	US-09-079-431B-6	Sequence 6, Appli
32	115.5	3.1	609	2	US-09-252-991A-32771	Sequence 23771, A
33	115	3.1	715	2	US-09-252-991A-32740	Sequence 32740, A
34	113.5	3.1	739	2	US-09-252-991A-32778	Sequence 32778, A
35	113.5	3.1	2509	1	US-08-469-005A-10	Sequence 10, Appl
36	112.5	3.0	1208	2	US-09-463-702A-2	Sequence 2, Appli
37	112.5	3.0	1208	2	US-09-699-135-2	Sequence 2, Appli
38	112.5	3.0	2418	2	US-09-949-016-10703	Sequence 10703, A
39	112.5	3.0	2511	2	US-09-361-907-2	Sequence 2, Appli
40	111.5	3.0	442	1	US-08-208-108-2	Sequence 2, Appli
41	110.5	3.0	683	2	US-09-949-016-11117	Sequence 11117, A
42	110.5	3.0	1147	1	US-08-131-365B-38	Sequence 38, Appl
43	110.5	3.0	1147	1	US-08-668-123-38	Sequence 38, Appl
44	110.5	3.0	1164	2	US-09-949-016-9845	Sequence 9845, Ap
45	110	3.0	386	2	US-09-605-703B-2514	Sequence 2514, Ap

ALIGNMENTS

RESULT 1

US-09-287-354-6
; Sequence 6, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mouse
US-09-287-354-6

Query Match		99.4%;	Score 3682;	DB 2;	Length 1182;
Best Local Similarity		99.4%;	Pred. No. 0;		
Matches 689;		Conservative 1;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKLEEDSSATSEGGGPGPEASLNKGL	60	
Db	490	VTQCSCVQAAGEVGLTGH	SQKRRSPLEEKLEEDSSATSEGGGPGPEASLNKGL	549	
QY	61	AKHLLSGLGDRLCRLLRKRE	ALAWAQREGQGPAMTDSFGIPHCCSRCHHGLFNTHWRC	120	
Db	550	AKHLLSGLGDRLCRLLRKRE	ALAWAQREGQGPAMTDSFGIPHCCSRCHHGLFNTHWRC	609	
QY	121	SHCSHRLCVACGRIAGACN	REKTSQEOHTDDCAQEAHAACSLILTOPVSSQALAE	180	
Db	610	SHCSHRLCVACGRIAGACN	REKTSQEOHTDDCAQEAHAACSLILTOPVSSQALAE	669	
QY	181	TYMHQAWAKFDIRGHCF	COVDARVWAPGDGGQKEPTKPTPTPPQPSGNSNRDKIKE	240	
Db	670	TYMHQAWAKFDIRGHCF	COVDARVWAPGDGGQKEPTKPTPTPPQPSGNSNRDKIKE	729	
QY	241	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT	300	
Db	730	ETPDSTESPAEDGAGRS	PLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT	789	
QY	301	NILDSTIAOVVERKIOE	KALGCLRAGSLPLSPVTRLSPPGALLWQSPRPK	360	
Db	790	NILDSTIAOVVERKIOE	KALGCLRAGSLPLSPVTRLSPPGALLWQSPRPK	849	
QY	361	HGFHFLQEHWRQCPVL	VSGIQKTLRLSLGWMEALGTLGGQVQSALTALGPPQPTNLDSTA	420	

Db 850 HGFLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPPQPTNLDSTA 909
QY 421 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 910 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLASSLPPEYCAHOGKLNLSYLP 969
QY 481 LGLTLHPLPOLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 970 LGLTLHPLPOLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 1029
QY 541 GLDGEGLWSPGSGTSTVWHVFRQAQDAQIRRFQWVCPAGAGTLEPGAPGSCYLDAGLRR 600
Db 1030 GLDGEGLWSPGSGTSTVWHVFRQAQDAQIRRFQWVCPAGAGTLEPGAPGSCYLDAGLRR 1089
QY 601 RLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAQLCHQG 660
Db 1090 RLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAQLCHQG 1149
QY 661 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 693
Db 1150 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 1182

RESULT 2
US-09-287-354-5
; Sequence 5, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: PRT
; ORGANISM: Rat
US-09-287-354-5

Query Match 94.8%; Score 3510; DB 2; Length 1207;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHSSQKSRSPLEEKLEEDSDATSSEGGGPGPEASLNKGL 60
Db 517 VTQCSCVQAAGEVEILTSHSQSKHKLPLEKPL-EEDSCATSEGGGS-SPEASLNKGL 574
QY 61 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 120
Db 575 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 634
QY 121 SHCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQVSSQALAE 180
Db 635 SHCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQVSSQALAE 694
QY 181 TVMHOAWAKFDIRGHCFQVDARVWAPDGGQOQKEPTKPTPQPCNGSDSNRTKDKE 240
Db 695 TVMHOAWAKFDIRGHCFQVDARVWAPDGGQOQKEPTKPTPQPCNGSDSNRTKDKE 754
QY 241 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 300
Db 755 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIT 814
QY 301 NILDSTIAQVVERKIQKALGPGLRAGSLRGLSLPLSPVTRTLPSPGALLWLQEPRPK 360
Db 815 NILDSTIAQVVERKIQKALGPGLRAGSLRGLSLPLSPVTRTLPSPGALLWLQEPRPK 874
QY 361 HGFLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQSLTALGPPQPTNLDSTA 420

Db 875 HGFLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQTLTALGPPQPTSLDSTA 934
QY 421 FWEGFHPETRPKLDGSGVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKLNLSYLP 480
Db 935 FWKGFSPHPEARPKLDGSGVLLHRTPLGDKDSESVENLASSLPPEYCAHOGKLNLSYLP 994
QY 481 LGLTLHPLPOLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 540
Db 995 LGLTLHPLPOLWAAAGVNSHRHGLGTKNLCVEVSDLSILVHAEAOQLPPWYRAQKDFLS 1054
QY 541 GLDGEGLWSPGSGTSTVWHVFRQAQDAQIRRFQWVCPAGAGTLEPGAPGSCYLDAGLRR 600
Db 1055 GLDGEGLWSPGSGTSTVWHVFRQAQDAQIRRFQWVCPAGAGTLEPGAPGSCYLDAGLRR 1114
QY 601 RLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAQLCHQG 660
Db 1115 RLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSQAQLCHQG 1174
QY 661 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 693
Db 1175 ASLPPDHRMLYAQMDDRAVFOAVKAAVGAALQEA 1207

RESULT 3
US-09-287-354-4
; Sequence 4, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-4

Query Match 83.6%; Score 3095.5; DB 2; Length 1189;
Best Local Similarity 83.2%; Pred. No. 1.7e-282;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHSSQKSRSPLEEKLEEDSDATSSEGGGPGPEASLNK 59
Db 494 LAQCSCQAAGGSGGHACHSQQVRSPLGGELQEEEDTATNSSSEGGPGSPDRSLSTG 553
QY 60 LAXHLLSGLDRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 119
Db 554 LAXHLLSGLDRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 613
QY 120 CSCHSHRLCVACGRITAGAKNREKTSQEQHTDDCAQAGHAACSLIITQVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSEBCEQAGHAACSLMLTQVSSQALAE 673
QY 180 STVMHOAWAKFDIRGHCFQVDARVWAPDGGQOQKEPTKPTPQPCNGSDSNRTKD 239
Db 674 STAHQVWAKFDIRGHCFQVDARVWAPDGGQOQKEPTKPTPQPCNGDTHRTKSIK 733
QY 240 ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDR 299
Db 734 ETPDSAETPAEDRAGRGPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDR 793
QY 300 TNILDSITIAQVVERKIQKALGPGLRAGSLRGLSLPLSPVTRTLPSPGALLWLQEP 358
Db 794 TNILDSITIAQVVERKIQKALGPGLRAGSLRGLSLPLSPVTRTLPSPGALLWLQEP 853
QY 359 -PKHGHFLFQEHWRQGPVLVSGIQKTLRLSLWGMEALGTLGGQVQSLTALGPPQPTNLD 417

Db 854 CPRGFHLFOEHWKQGPVLVSGIQTLOGNMGTEALGALGGQVQALSPLGPPQPSLSG 913
QY 418 STAFWEGFSPETPKLDEGSVLLHRTLGDKASRVQNLVSSILPLPEYCAHQCKNLAS 477
Db 914 STTFWEGFSPPELKPCKDEGSVLLHRTLGDETSRVENLAASLPLPEYCALHCKNLAS 973
QY 478 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDLSILVHAEALPMPMYRAQKD 537
Db 974 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVADLVSLVHADTLPMPMYRAQKD 1033
QY 538 FLGLDEGLWSPGSGTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPCAGPSCYLDAG 597
Db 1034 FLGLDEGLWSPGSGTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPCAGPSCYLDAG 1093
QY 598 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSALC 657
Db 1094 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSALC 1153
QY 658 HQGASLPDPHMLYAQMDRAVFOAVKAAVGTQLEAK 693
Db 1154 HQGASLPDPHMLYAQMDRAVFOAVKAAVGTQLEAK 1189

RESULT 4

US-09-949-016-6931
; Sequence 6931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6931
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6931

Query Match 83.6%; Score 3095.5; DB 2; Length 1189;
Best Local Similarity 83.2%; Pred. No. 1.7e-282;
Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHKSRRSPLEKQLEEDSSATSEGGGPGPEASLNGK 59
Db 494 LAQCSCQAAGGEGGHACHSQVRRSPGLGELQEEEDTATNSSEEGPGSGDRLSTG 553
QY 60 LAKHLLSGLDRLCLLRKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 119
Db 554 LAKHLLSGLDRLCLLRKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRIAGAGNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLILTOFVSSQALAE 673
QY 180 STVHQAWAKFDIRGHCFQVDARVWAPGGGQOKEPTKPTPTPQSCNGDSNRKDK 239
Db 674 STAHQVWVWFDIRGHCFQVDARVWAPGGGQOKEPTKPTPTPQSCNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
Db 734 EETPDSTESPAEDGAGSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPVRTRLSPPGALLWLQEP - 358

Db 794 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPVRTRLSPPGALLWLQEP - 853
QY 359 -PKHGFHLFOEHWKQGPVLVSGIQTLOGNMGTEALGALGGQVQALSPLGPPQPSLSG 417
Db 854 CPRGFHLFOEHWKQGPVLVSGIQTLOGNMGTEALGALGGQVQALSPLGPPQPSLSG 913
QY 418 STAFWEGFSPETPKLDEGSVLLHRTLGDKASRVQNLVSSILPLPEYCAHQCKNLAS 477
Db 914 STTFWEGFSPPELKPCKDEGSVLLHRTLGDETSRVENLAASLPLPEYCALHCKNLAS 973
QY 478 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVSDLSILVHAEALPMPMYRAQKD 537
Db 974 YLPGLGLTLHPLEPOLMAAYGVNSHRHGLTKNLCVEVADLVSLVHADTLPMPMYRAQKD 1033
QY 538 FLGLDEGLWSPGSGTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPCAGPSCYLDAG 597
Db 1034 FLGLDEGLWSPGSGTSTVHVFRAQDAQRIIRFLQWVCPAGAGTLEPCAGPSCYLDAG 1093
QY 598 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSALC 657
Db 1094 LRRRLREMGVSCWTLLOAPGEAVLVPAGAPHQVQGLVSTISVTOHFLSPETSALSALC 1153
QY 658 HQGASLPDPHMLYAQMDRAVFOAVKAAVGTQLEAK 693
Db 1154 HQGASLPDPHMLYAQMDRAVFOAVKAAVGTQLEAK 1189

RESULT 5

US-09-287-354-2
; Sequence 2, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-2

Query Match 83.4%; Score 3088.5; DB 2; Length 984;
Best Local Similarity 83.0%; Pred. No. 5.9e-282;
Matches 578; Conservative 40; Mismatches 75; Indels 3; Gaps 2;
QY 1 VTQCSCVQAAGEVGLTGHKSRRSPLEKQLEEDSSATSEGGGPGPEASLNGK 59
Db 289 LAQCSCQAAGGEGGHACHSQVRRSPGLGELQEEEDTATNSSEEGPGSGDRLSTG 348
QY 60 LAKHLLSGLDRLCLLRKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 119
Db 349 LAKHLLSGLDRLCLLRKEREALAWAQREGQGPANTEPGPHCCSRCHGLFNTHWR 408
QY 120 CSHCSHRLCVACGRIAGAGNREKTSQEQHTDDCAQEAAGAACSLILTOFVSSQALAE 179
Db 409 CPRCSHRLCVACGRVAGTGRAREKAGFQEQSAEECTQEAAGAACSLILTOFVSSQALAE 468
QY 180 STVHQAWAKFDIRGHCFQVDARVWAPGGGQOKEPTKPTPTPQSCNGDSNRKDK 239
Db 469 STAHQVWVWFDIRGHCFQVDARVWAPGGGQOKEPTKPTPTPQSCNGDTHRTKSIK 528
QY 240 EETPDSTESPAEDGAGSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
Db 529 EETPDSTESPAEDGAGSPLPCPSLCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 588
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGSLRKLGLSLPVRTRLSPPGALLWLQEP - 358

Db 589 TNLDSIIAQVVERKIQEALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP 648
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWMEALGTGGQVQSLTALGPPQPTNLD 417
Db 649 CPRRGHFLFQEHWRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLGPPQSSLG 708
QY 418 STAFWEGFHPETRPKLDGEGSVLLHRTLGDKDASRVQNLVSSLPLPEYCAHOGKUNLAS 477
Db 709 STTFWEGFSPWELRPKSDGEGSVLLHRTLGDKDASRVQNLVSSLPLPEYCAHOGKUNLAS 768
QY 478 YLPLGLTLHLEPQLWAAYGVNSHRHGLGTKNLCVEVSDLI SILVHAEALPWPYRAQKD 537
Db 769 YLPPGLALRLEPQLWAAYGVNSHRHGLGTKNLCVEVADLVSLVHARTPLPAWHAQKD 828
QY 538 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSCYLDAG 597
Db 829 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSCYLDAG 888
QY 598 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 657
Db 889 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 948
QY 658 HQGASLPPDHRMLYAQMDRAVFOAKAAGALQEA 693
Db 949 HOGPSLPPDCHLLYAQMDWAVFOAKVAVGTLOEAK 984

RESULT 6

US-09-287-354-3

; Sequence 3, Application US/09287354
; Patent No. 6348348
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: THOMPSON-20263/0243435
; CURRENT APPLICATION NUMBER: US/09/287,354
; CURRENT FILING DATE: 1999-04-07
; EARLIER APPLICATION NUMBER: US 60/080,888
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-287-354-3

Query Match 83.1%; Score 3077.5; DB 2; Length 1189;
Best Local Similarity 82.9%; Pred. No. 8.6e-281;
Matches 577; Conservative 37; Mismatches 79; Indels 3; Gaps 2;

QY 1 VTQCSCVQAAGEVGLTGHSHQSRRLPL-EEKOLEEEDSSATSEEGGGPGPEASLNGK 59
Db 494 LAQCSCQAAGEGGGHACHSQVRRPLGLGELQOEDTATNSSSEGGPGSPRLSTG 553
QY 60 LAKHLLSLGLDRLCLRLKEREALAWAQREGQGPAMTSDSPGPHCCSRCHHGLFNTHWR 119
Db 554 LAKHLLSLGLDRLCLRLGEREALAWAQREGSQGPAMTSDSPGPRCCSRCHHGLFNTHWR 613
QY 120 CSHCSHRLCVACGRIAGAKNREKTSQEQHTDDCAQEAHAACSLILTQVSSQALAE 179
Db 614 CPRCSHRLCVACGRVAGTRAREKAGFQEQSAEECTQEAHAACSLMLTQVSSQALAE 673
QY 180 STVMHQAQKFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPQPSGNGDSNRTKDIK 239
Db 674 STAMQVWVXFDIRGHCFQVDARVWAPGDGQKQKTEPTPTPQPSGNGDTHRTKSIK 733
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 299
Db 734 EETPDSAETPAEDRAGRPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRI 793
QY 300 TNLDSIIAQVVERKIQEALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP - 358

Db 794 TNLDSIIAQVVERKIQEALGPGLRAGLGLPLSPVPRPLPPPGALLWLQEPQP 853
QY 359 -PKHGFHLFQEHWRQGPVLVSGIQKTLRLSLWMEALGTGGQVQSLTALGPPQPTNLD 417
Db 854 CPRRGHFLFQEHWRQGPVLVSGIORTLQGNLWGTALGALGGQVQALSPLAPPQSSLG 913
QY 418 STAFWEGFHPETRPKLDGEGSVLLHRTLGDKDASRVQNLVSSLPLPEYCAHOGKUNLAS 477
Db 914 STTFWEGFSPWELRPKSDGEGSVLLHRTLGDKDASRVQNLVSSLPLPEYCAHOGKUNLAS 973
QY 478 YLPLGLTLHLEPQLWAAYGVNSHRHGLGTKNLCVEVSDLI SILVHAEALPWPYRAQKD 537
Db 974 YLPPGLALRLEPQLWAAYGVNSHRHGLGTKNLCVEVADLVSLVHARTPLPAWHAQKD 1033
QY 538 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSCYLDAG 597
Db 1034 FLSGLDGEGLSWPGSGTSTVWHVFRADQRIIRRFLOMVCPCAGAGTLEPGAGSCYLDAG 1093
QY 598 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 657
Db 1094 LRRRLREEWGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTHFLSPETSALSALC 1153
QY 658 HQGASLPPDHRMLYAQMDRAVFOAKAAGALQEA 693
Db 1154 HOGASLPPDCHLLYAQMDWAVFOAKVAVGTLOEAK 1189

RESULT 7

US-09-949-016-7891
; Sequence 7891, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7891
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7891

Query Match 63.0%; Score 2333.5; DB 2; Length 943;
Best Local Similarity 77.2%; Pred. No. 9.4e-211;
Matches 438; Conservative 35; Mismatches 63; Indels 31; Gaps 4;

QY 42 TSEGGGGPGPEASLNKGLAKHLLSLGLDRLCLRLKEREALAWAQREGQGPAMTSDSPG 101
Db 387 TAEQAGGQVEVDTSIG-NKVDVDSQHD-----EQGGQGPAMTSDSPG 429
QY 102 IPHCCSRCHHGLFNTHWRCSHRLCVACGRITAGAKNREKTSQEQHTDDCAQEAHA 161
Db 430 IPRCCSRCHHGLFNTHWRCPRCSHRLCVACGRVAGTRAREKAGFQEQSAEECTQEAHA 489
QY 162 ACSLILTFVSSQALAEALSTVMEHQAQKFDIRGHCFQVDARVWAPGDGQKQKTEPTKTP 221
Db 490 ACSLMLTFVSSQALAEALSTAMQVWVXFDIRGHCFQVDARVWAPGDGQKQKTEPTKTP 549
QY 222 PTPQPSGNGDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDR 281
Db 550 PTPQPSGNGDTHRTKSIKEETPDSDAETPADRAGRPLPCPSLCELLASTAVKLCIGHDR 609

Qy	282	IHWAFAPVTPALPSDDRIITNILDIIIAQWVERKIQEALGPGCLRAGSLRKGLSLPSPV	341
Db	610	IHWAFAPVTPALPSDDRIITNILDIIIAQWVERKIQEALGPGCLRAGSLRKGLSLPSPV	669
Qy	342	RTRLSPFGALLMLQEPFR--PKHGPHLFQEHWRQGPVLVSGIQKTLRLSLWGMGMEALCTLG	399
Db	670	RRLPLPPGALLMLQEPQPCPRRGPHLFQEHWRQGPVLVSGIQRTLQGNLWGTEALGALG	729
Qy	400	GOVOSITALGPPOPTNLDSTAFWEGFSHPETPKLDEGSVLLIHRITLGDKDA	452
Db	730	GOVALUSPLGPPQPSLSGTTFWEGFSHPETPKLDEGSVLLIHRALGDSDTLRGLLGM	789
Qy	453	-----RVQNLVSLPLPEYCAHQKGLNLASYLPLGLTLHPLEPOLWAAYGVNSHRGHG	507
Db	790	TPYPRDVENLAASLPLPEYCALHGKGLNLASYLPPCLALRPLEPOLWAAYGVSPIRHHG	849
Qy	508	KNLCEVRSDLIITLVHAEAQLPWYRAQKDFLSGLDCEGLWSPGOSTVWHVFRAQDAQ	567
Db	850	KNLCEVADLVILVHADTPLPAWHRAQKDFLSGLDCEGLWSPGQSVSTVWHVFRAQDAQ	909
Qy	568	RIRRFLOWVCPAGAGTLBPAGPGCYL	594
Db	910	RIRRFLOWVCPAGAGALBPAGPGCYL	936

RESULT 8

US-09-640-211A-1138
; Sequence 1138, Application US/09640211A

; Patent No. 6833446
; GENERAL INFORMATION:

APPLICANT: Wood, Marion
APPLICANT: Shenk, Michael A.
APPLICANT: McGrath, Annette
APPLICANT: Glenn, Matthew

INVENTOR: STEVEN J. MATHIAS

TITLE OF INVENTION: Compositions and Methods for the

TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

CURRENT FILING DATE: 2000-08-16

NUMBER OF SEQ

; SOFTWARE: Fast

; SEO ID NO 1138

: LENGTH: 177

TYPE: PRT

LIFE: FRI
ORGANISM: Pinus radiata

ORGANISM: *Pinus taeda*
 IIS-09-640-211A-1138

RESULT 9

US-09-514-247A-10

03-03 214-217A-10 : Sequence 10. Application US/09514247A

: Sequence ID, APPLIC
: Patent No. 6365361; FACILE NO. 8363361
; GENERAL INFORMATION:
; PRESENT TENSE C

; APPLICANT: TANABE SEIYAKU CO. LTD.
 ; APPLICANT: TANIGUCHI, Tomoyasu
 ; APPLICANT: MIZUKAMI, Junko
 ; TITLE OF INVENTION: METHOD FOR IDE

; FILE REFERENCE: TANIGUCHI=6

; CURRENT APPLICATION NUMBER: US/09/514,247A

; CURRENT FILING DATE: 2000-02-28

; PRIOR APPLICATION NUMBER: PCT/JP98/03734

Db 2315 SPOQHMLSGQPQASHLPQG 2333

RESULT 10

US-09-538-092-1370

; Sequence 1370, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:

; APPLICANT: Giot, Loic

; APPLICANT: Mansfield, Traci A.

; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

; FILE REFERENCE: 15966-542

; CURRENT APPLICATION NUMBER: US/09/538,092

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: 60/127,352

; PRIOR FILING DATE: 1999-04-01

; PRIOR APPLICATION NUMBER: 60/178,965

; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387

; SOFTWARE: CuratPatSeqFormat Version 0.9

; SEQ ID NO 1370

; LENGTH: 2442

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (0)...(0)

; OTHER INFORMATION: Polypeptide Accession Number Q92793

US-09-538-092-1370

Query Match 4.1%; Score 152.5; DB 2; Length 2442;

Best Local Similarity 18.2%; Pred. No. 0.00024;

Matches 156; Conservative 104; Mismatches 282; Indels 317; Gaps 42;

Qy 30 EEKQLEEDSSATSEEGGGPG-----PEASLNKGLAKHLLSLGL 69

Db 1559 EEEERKEEETAASETTGSGDSGNKAKNNKNTNKNKSSIRANKKPKSPMENVNDLS 1618

Qy 70 DRLCLRLKERE-----LAWAQREGQGPAM----- 95

Db 1619 QKLYATMEKHEVFFVIHLHAGPVINTLPVDPDPLSCDLMDGRDAFLTLARDKHWEF 1678

Qy 96 -----TEDSGIPHCRCRHGLFNTHWRSCH-SHRLCVAC----- 131

Db 1679 SSLRRSKWSTLCMLVELHTQODRFVYTCNECKHV-ETRWHTVCBDYDLINCYNKTS 1737

Qy 132 -----GRIAGAGNREKTGSGEHTDCAEAGHAA-----CSLIITQF 170

Db 1738 HAHKWVKGGLDDEGSSQGEPSQSKPQESRRVRSIQICISLVHACOCRNANCSL----- 1792

Qy 171 VSSQALAEISTVMHQAWEKDIRG-----HC-----FC----- 198

Db 1793 PSCQKMRV-VQHTGCKRTNGGCPVKQOLIALCCYHAKHCQENKCPVPFCLNIKHL 1850

Qy 199 ----QVDARVWAPGGGOKE-----PTEKTP--PTPQSPCNGDS 232

Db 1851 RQOQIQHRL-----QQAQLMRRRMATNRNVPOQSLPSPTAPPCTPTQPS----- 1898

Qy 233 NRTKDIKEETPDSTESPAEDGAGSPPLPCSLCELLASTAVKLCGLGHDRIHMAFVTPA 292

Db 1899 -----TFQTPQPPAQPPSPVSPSPAGFPFVSARTQPTTSTGKTSQVP---APPPPA 1949

Qy 293 LPSSDRTITNLDSTIAQVVERKIQEK-----ALGPGLRAGSGLRKGLSLPLS--- 339

Db 1950 QPPAAVE-----AAQIEREAQOQOHLRYVNNINMPPG-FRTGMPGSGMAVPSLVN 2002

Qy 340 -----PVRTRISPGCALLWLQEPKPGHFLFQEHWRQGPVLVSGIKTLURLSLMG 391

Db 2003 PRPNQVSGVPMSM-PPCQ--WQAPLP-----QQQP--MPGLPR-----PVIS 2041

Qy 392 MEALGTGGQVQSLTALGPPQ---PTNL-DSTAWEFGFSPHETPK---LDEGSVLLHR 444

Db 2042 MQAQAAG--PRMPSVQPPRSISPSALQDLRLTKGPSQQQQQVNLNLSKPMQMAA 2099

Qy 445 TLGDKDASRVQNLVSSSLPLPEYCAHOG-----KLNLSYPLPLGLTLHPLPQLW 493

Db 2100 FIKQRTAKYVANQPGMQPQGLSQPGMQPQPMHQPSLQNLNMAQAGVPRPGVPPQQ 2159

Qy 494 AAYGVN-----SHRGHLGTKNLCVEVSDLI--SILVHAEALPFWYRAQKD--- 537

Db 2160 AMGGLNPOGALNINPCHNPNMNSN--PQYREMLRROLLQOQOQOQOQOQOQOQOQ 2217

Qy 538 ---FLSGLDEGLW-----SPGSGTSTVMHVHFAQAQIRIRFL-----QMVCPAG-- 580

Db 2218 SAGMAGMAGHGFQPGQPGGYPAM-----QQQRMQOHLPLQSSMGMAAQMGQL 2271

Qy 581 --AGTLEPGAPGSCYLDAGLRRLREWGVSCWTLLOAPGEAVLVAGAPHQVQGLVSTI 638

Db 2272 GQMGPGLGADSTPNITQALQORILQQ-----QQMKQIQIGSP---GQPNPM 2314

Qy 639 SVTQHLS--PETSALSAQ 655

Db 2315 SPOQHMLSGQPQASHLPQG 2333

RESULT 11

US-08-227-536-2

; Sequence 2, Application US/08227536

; Patent No. 5658784

; GENERAL INFORMATION:

; APPLICANT: Eckner, Richard

; APPLICANT: Ewen, Mark

; APPLICANT: Livingston, David

; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION

; TITLE OF INVENTION: FACTOR P300 AND USES OF P300

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes

; STREET: Ten Post Office Square

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,536

; FILING DATE: 14-APR-1994

; CLASSIFICATION: 436

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams Ph.D., Kathleen A.

; REGISTRATION NUMBER: 34,380

; REFERENCE/DOCKET NUMBER: DFCI-308XX

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-2290

; TELEFAX: (617) 451-0313

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2414 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-227-536-2

Query Match 4.0%; Score 146.5; DB 1; Length 2414;

Best Local Similarity 19.0%; Pred. No. 0.00089;

Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;

Qy 29 LEE--KOLEEDSSATSEEGGGGPGPE-----ASLNKGLAK--- 62

Db 1513 LEESIKLEQEEERKEENTSNESTDVTKGDSKNAKNNKNTKSNKSSLSRGNKKPG 1572

Qy 63 --HLLSGLGRLCLRLKEREALAWAQREGQPA-----MTEDSPCIP----- 103

```
Db 1573 MPNVSNLSQKYATMEKHEVF-FVIRLIAGPAANSPPVDPDPLIPCDLMDGRDAFL 1631
QY 104 -----HCCSRCHHGLFNTHWRSCHC-SHR 126
Db 1632 TLARDKHLEFSSLRRAQWSTMCMVLVELHTQSDRFFVTTCNECKHV-ETRWHTVCEYD 1690
QY 127 LCVACGRIAGAGNREKTG-----SQEQHTDDCAQEAAGHAA----- 162
Db 1691 LCITCYNTKNHDHMEKGLGLDDESNQAAATQSPGDSRRLSIQRCISLVHACQCRN 1750
QY 163 --CSLILTQVSSQALAEELSTVHQAWAKFDIGHC-----FCQVD----- 201
Db 1751 ANCSL-----PSCQKMKRV--VQHTGCKRKTNGGCPICKQIALCCYHAKHCQENKCPV 1803
QY 202 -----ARVWAPGDGQOK-----BPEKTPPTP-----OP 226
Db 1804 PFCNLTKQKLRQOOLQHRLOQAQMLRRRMAWMQRTGVVQOQGLPSPPTPATPTTPTGQQP 1863
QY 227 SCNGDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAF 286
Db 1864 T-----TPQTPQTSQP-----QTPPNMSM----- 1883
QY 287 APVTPALPSDDRTNILDIIAQVVERKIQEAKGGLRAGSLRGKLSLPLSPVTRLS 346
Db 1884 ---PPYLP-----RTQAAGP-----VSQGAAGQVTPPTPTPTAOP 1916
QY 347 P-PGALLWLQERPP---KHGPHLFQEHWRQGPVLSGQTKTLRLSLWGMEALGTGGQV 402
Db 1917 PLPG-----PPPTAVEMAMQIQRAAETQORNAHVQIFORPIOHQMPMPMPAPMG--- 1966
QY 403 QSLTALGPPQPTNLDSTAFWEGF--SHPETRPKLDGEGSVLLHRTLGLDKDASRVQNLVSS 460
Db 1967 -----KNPPTMTRGSPGHLEPGMGPTCMQOQPPWSQGL-----PQOQLOSG 2009
QY 461 LPLP---EYCAHQGKLNILASYPLG-LTLHPLPQLWAAAYGVNSHRHGLTKNLCVEVSD 516
Db 2010 MPRPAMMSVAHQGPLNMAPQGLGQVIGISPLKPGTVSQALQNLRLTLRSPSSPLQOQ 2069
QY 517 LISILVHAEALPPWY---RAOK-----DFLSGLDGEGLWSPGSGTSTVWVFRQDAQ 567
Db 2070 VLSIL-HANPQLLAAFIKORAAKYANSNPQIPFGQPMGQGLPQPTM-----PQOQGV 2124
QY 568 RIRFLQWVCPAGAGTLEPGAP 589
Db 2125 HSNPAMQNNPMQAGVQVQAGLP 2146
```

RESULT 12

```
US-09-538-092-1289
; Sequence 1289, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapacSeqformatter Version 0.9
; SEQ ID NO 1289
; LENGTH: 2414
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q09472
```

US-09-538-092-1289

```
Query Match 4.0%; Score 146.5; DB 2; Length 2414;
Best Local Similarity 19.0%; Pred.No.0.00089;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;
QY 29 LEE--KQLEEDSSATSEEGGGPGPE-----ASLNKGLAK--- 62
Db 1513 LEEISIKELQEBEERKREENTSNESDVTKGDSKNAKXNNKKTYSKNKSLSRGNKKKPG 1572
QY 63 --HLSGLGDRCLRLRKEREALAWAQREGQPA-----MTEDSGIP----- 103
Db 1573 MPNVSNLSQKYATMEKHEVF-FVIRLIAGPAANSPPVDPDPLIPCDLMDGRDAFL 1631
QY 104 -----HCCSRCHHGLFNTHWRSCHC-SHR 126
Db 1632 TLARDKHLEFSSLRRAQWSTMCMVLVELHTQSDRFFVTTCNECKHV-ETRWHTVCEYD 1690
QY 127 LCVACGRIAGAGNREKTG-----SQEQHTDDCAQEAAGHAA----- 162
Db 1691 LCITCYNTKNHDHMEKGLGLDDESNQAAATQSPGDSRRLSIQRCISLVHACQCRN 1750
QY 163 --CSLILTQVSSQALAEELSTVHQAWAKFDIGHC-----FCQVD----- 201
Db 1751 ANCSL-----PSCQKMKRV--VQHTGCKRKTNGGCPICKQIALCCYHAKHCQENKCPV 1803
QY 202 -----ARVWAPGDGQOK-----BPEKTPPTP-----OP 226
Db 1804 PFCNLTKQKLRQOOLQHRLOQAQMLRRRMAWMQRTGVVQOQGLPSPPTPATPTTPTGQQP 1863
QY 227 SCNGDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAF 286
Db 1864 T-----TPQTPQTSQP-----QTPPNMSM----- 1883
QY 287 APVTPALPSDDRTNILDIIAQVVERKIQEAKGGLRAGSLRGKLSLPLSPVTRLS 346
Db 1884 ---PPYLP-----RTQAAGP-----VSQGAAGQVTPPTPTPTAOP 1916
QY 347 P-PGALLWLQERPP---KHGPHLFQEHWRQGPVLSGQTKTLRLSLWGMEALGTGGQV 402
Db 1917 PLPG-----PPPTAVEMAMQIQRAAETQORNAHVQIFORPIOHQMPMPMPAPMG--- 1966
QY 403 QSLTALGPPQPTNLDSTAFWEGF--SHPETRPKLDGEGSVLLHRTLGLDKDASRVQNLVSS 460
Db 1967 -----KNPPTMTRGSPGHLEPGMGPTCMQOQPPWSQGL-----PQOQLOSG 2009
QY 461 LPLP---EYCAHQGKLNILASYPLG-LTLHPLPQLWAAAYGVNSHRHGLTKNLCVEVSD 516
Db 2010 MPRPAMMSVAHQGPLNMAPQGLGQVIGISPLKPGTVSQALQNLRLTLRSPSSPLQOQ 2069
QY 517 LISILVHAEALPPWY---RAOK-----DFLSGLDGEGLWSPGSGTSTVWVFRQDAQ 567
Db 2070 VLSIL-HANPQLLAAFIKORAAKYANSNPQIPFGQPMGQGLPQPTM-----PQOQGV 2124
QY 568 RIRFLQWVCPAGAGTLEPGAP 589
Db 2125 HSNPAMQNNPMQAGVQVQAGLP 2146
```

RESULT 13

```
PCT-US95-04682-2
; Sequence 2, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
```

461 LPLP---EYCAHQGKLNLSVLPGL-LTLHPLPQLMAAYGVNSHRHGLTGKNCULCEVSD 516
2010 MPRPAMSVAGHQQPLNMAPOGLGQGVISPLKPTVTSQOALQNLRLTLRSFSSPLOQQQ 2069
517 LISILVHAEALPPWY---RAQK-----DFLSGLDGEGLWSPSGTSTVWHVFRQADAO 567
2070 VLSIL-HANPQLAFAIKORAAKYANSNPQIPGQPGMPQGGPGLQPTM-----PQOQGV 2124
568 RIRFLQWVCPAGAGTLEPGAP 589
2125 HSNPAMQNNPMQAGVQRAGLP 2146
RESULT 14
US-08-194-468-2
Sequence 2, Application US/08194468
Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
TITLE OF INVENTION: RESPONSIVE GENES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-468-2
Query Match 3.9%; Score 145.5; DB 1; Length 2441;
Best Local Similarity 18.8%; Pred. No. 0.0011;
Matches 158; Conservative 101; Mismatches 290; Indels 301; Gaps 42;
30 EEKQLEEDSSATSEGGGPGP-----PEASLNKGLAKHLGLG 69
1560 EESERKKEETAASEITEPGSGDSKNAKNNKNTKNKSSISRAKKKPSMPNVSNDS 1619
70 DRCLRLRKERREALAWAQREGQGPATPESPGIP----- 103
1620 QKLYATMEKHKEVF-FVIRLIAGPAANSPLPVPDPLPCDLMDGRDAFLTLARDKHE 1678
104 -----HCCSRCHHGLFNTHWRCSC-HSRLCVAC----- 131
1679 FSLRRSKWSTLCMLVELHTQGGDRFVITCNECKHV-ETRWHTCTVEDYDLINCYNK 1737
132 -----GRIAGAGKNREKTSQEQHDTDDCAQEAAGHAA-----CSLILQ 169
1738 SHTHKVMKWLGLDDEGSSQGEPSQSPQBSRRLSIQRCSLVLVHACQCRNANCSL----- 1793

408; Score 146.5; DB 4; Length 2414;
Best Local Similarity 19.0%; Pred. No. 0.00089;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;
29 LEB--KOLEEDSSATSEGGGPGP-----ASLNKGLAK--- 62
1513 LEBSEKELEGEERKEENTSNSTVTYKGDKNVAKNNKNTSKNKSLSRGNKKKPG 1572
63 --HLLSGLDRLCLRLKEREALAWAQREGGPA-----MTEDSPGIP----- 103
1573 MPNVSNDSQKLYATMEKHKEVF-FVIRLIAGPAANSPLPVPDPLPCDLMDGRDAFL 1631
104 -----HCCSRCHHGLFNTHWRCSC-HSRLCVAC----- 131
1632 TLARDKHLEFSSLRRAQWSTMCMLVELHTQGGDRFVITCNECKHV-ETRWHTCTVEDYD 1690
127 LCVACGRIAGKUREKTG-----SOEQHTDDCAQEAAGHAA----- 162
1691 LCITCYNTKNDHKWKEKLGLDDESNNQQAATQSPGDSRRLSIQRCSLVLVHACQCRN 1750
163 --CSLILQTVSSQALAEISTVMHQAQWAKFDIRGHC-----FCQVD----- 201
1751 ANCSL-----PSCQMKRV--VQHTKGCKRTNGGCPICKQLIALCCVHAKQCNKCPV 1803
202 -----ARWAPGCGQOK-----EPEKTPPTP---OP 226
1804 PFCLNIRKQLRQQLQHLRQQAQMLRRRMAQMORTGVVQGGQLPSPTATPTTPTGQOP 1863
227 SCNGDSNRKTDIKEETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCGLHDIRHMAF 286
1864 T-----TQQTQFTSQP-----QTPPNSM----- 1883
287 APVTPALPSDDRITNILDSTIAQVVERKIQEALPGRLAGSGLRKGLSLPLSPVTRLS 346
1884 ---PPYLP-----RTQAAGP---VSQKAAGQVTPPTPPTQAQ 1916
347 P-PCALLWQPRP---KHGPHLFOEHWRCQPVLVSGIQKTLRLSLMGWALGTILGQV 402
1917 PLPG-----PPPTAVEMAMQIQRAAETQRMARVQIFQPIQHQPMPPTFMAPMG--- 1966
403 QSLTALGPPQPTNLDSATFMEGF--SHPETPKLDEGSVLLHRLTLGDKASRVQNLVSS 460
1967 -----MNPPTMRGSPSHLEFGMGFTGNQOQPPNSQGL-----PQPQQLSQ 2009

Query Match 4.0%; Score 146.5; DB 4; Length 2414;
Best Local Similarity 19.0%; Pred. No. 0.00089;
Matches 141; Conservative 70; Mismatches 242; Indels 289; Gaps 33;
29 LEB--KOLEEDSSATSEGGGPGP-----ASLNKGLAK--- 62
1513 LEBSEKELEGEERKEENTSNSTVTYKGDKNVAKNNKNTSKNKSLSRGNKKKPG 1572
63 --HLLSGLDRLCLRLKEREALAWAQREGGPA-----MTEDSPGIP----- 103
1573 MPNVSNDSQKLYATMEKHKEVF-FVIRLIAGPAANSPLPVPDPLPCDLMDGRDAFL 1631
104 -----HCCSRCHHGLFNTHWRCSC-HSRLCVAC----- 131
1632 TLARDKHLEFSSLRRAQWSTMCMLVELHTQGGDRFVITCNECKHV-ETRWHTCTVEDYD 1690
127 LCVACGRIAGKUREKTG-----SOEQHTDDCAQEAAGHAA----- 162
1691 LCITCYNTKNDHKWKEKLGLDDESNNQQAATQSPGDSRRLSIQRCSLVLVHACQCRN 1750
163 --CSLILQTVSSQALAEISTVMHQAQWAKFDIRGHC-----FCQVD----- 201
1751 ANCSL-----PSCQMKRV--VQHTKGCKRTNGGCPICKQLIALCCVHAKQCNKCPV 1803
202 -----ARWAPGCGQOK-----EPEKTPPTP---OP 226
1804 PFCLNIRKQLRQQLQHLRQQAQMLRRRMAQMORTGVVQGGQLPSPTATPTTPTGQOP 1863
227 SCNGDSNRKTDIKEETPDSTESPAEDGAGRSPLPCPSLCCELLASTAVKLCGLHDIRHMAF 286
1864 T-----TQQTQFTSQP-----QTPPNSM----- 1883
287 APVTPALPSDDRITNILDSTIAQVVERKIQEALPGRLAGSGLRKGLSLPLSPVTRLS 346
1884 ---PPYLP-----RTQAAGP---VSQKAAGQVTPPTPPTQAQ 1916
347 P-PCALLWQPRP---KHGPHLFOEHWRCQPVLVSGIQKTLRLSLMGWALGTILGQV 402
1917 PLPG-----PPPTAVEMAMQIQRAAETQRMARVQIFQPIQHQPMPPTFMAPMG--- 1966
403 QSLTALGPPQPTNLDSATFMEGF--SHPETPKLDEGSVLLHRLTLGDKASRVQNLVSS 460
1967 -----MNPPTMRGSPSHLEFGMGFTGNQOQPPNSQGL-----PQPQQLSQ 2009

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-April-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xq999
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2414 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-04682-2

This Page Blank (Page 4)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:06:54 ; Search time 10 Seconds
(without alignments)
210.046 Million cell updates/sec

Title: US-10-712-629B-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFQAVKAAGALQEA 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_AA_New:*
- 1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pbp.*
 - 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pbp.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pbp.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pbp.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pbp.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pbp.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pbp.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	3.1	667	1 US-10-821-234-1477	Sequence 1477, Ap
2	102	2.8	1402	1 US-10-971-982-2	Sequence 2, Appli
3	101	2.7	447	1 US-10-967-527A-14	Sequence 14, Appl
4	97.5	2.6	658	1 US-10-821-234-921	Sequence 921, App
5	97	2.6	614	7 US-11-015-546A-20	Sequence 20, Appl
6	95.5	2.6	1184	1 US-10-131-826A-412	Sequence 412, App
7	95	2.6	4419	1 US-10-821-234-1155	Sequence 1155, Ap
8	93.5	2.5	595	7 US-11-182-946-9	Sequence 9, Appli
9	93	2.5	483	7 US-11-184-574-4	Sequence 4, Appli
10	93	2.5	915	1 US-10-821-234-1514	Sequence 1514, Ap
11	90.5	2.4	542	7 US-11-152-747-6	Sequence 6, Appli
12	90	2.4	777	1 US-10-821-234-1658	Sequence 1658, Ap
13	89.5	2.4	1618	1 US-10-984-645-2	Sequence 2, Appli
14	89	2.4	415	7 US-11-182-946-6	Sequence 6, Appli
15	88.5	2.4	419	1 US-10-821-234-1556	Sequence 1556, Ap
16	88.5	2.4	1138	7 US-11-012-762-28	Sequence 28, Appl
17	88	2.4	1694	7 US-11-135-855-36	Sequence 36, Appl
18	88	2.4	1709	7 US-11-135-855-35	Sequence 35, Appl
19	86.5	2.3	308	1 US-10-967-527A-30	Sequence 30, Appl
20	86.5	2.3	434	7 US-11-135-855-40	Sequence 40, Appl
21	86.5	2.3	653	1 US-10-131-826A-438	Sequence 438, App
22	86.5	2.3	653	7 US-11-135-855-25	Sequence 25, Appl
23	86	2.3	408	7 US-11-135-855-39	Sequence 39, Appl
24	86	2.3	1199	1 US-10-821-234-1126	Sequence 1126, Ap
25	85.5	2.3	1560	7 US-11-059-982-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-10-821-234-1477

; Sequence 1477, Application US/10821234

; Publication No. US20050255114A1

; GENERAL INFORMATION:

; APPLICANT: Labat, Ivan

; APPLICANT: Stache-Crain, Birgit

; APPLICANT: Andarmani, Suean

; APPLICANT: Tang, Y. Tom

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia

; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234

; CURRENT FILING DATE: 2004-04-07

; PRIOR APPLICATION NUMBER: US 60/462,047

; PRIOR FILING DATE: 2003-04-07

; NUMBER OF SEQ ID NOS: 1704

; SOFTWARE: pt_seq_genes Version 1.0

; SEQ ID NO 1477

; LENGTH: 667

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-821-234-1477

Query Match	3.1%	Score 115;	DB 1;	Length 667;
Best Local Similarity	20.4%	Pred. No. 0.046;		
Matches 112;	Conservative 55;	Mismatches 191;	Indels 192;	Gaps 25;
QY	2	TQCSCVQAAGE--VGVLT-----GHSQKRRS-----PLEEKQLEEDSSA-----	41	
DB	60	TKPQDMISAGGESVAGITAIAGKPGDKKKEKSLTPAVFVESPDPKPSGSGMDAALDDL	119	
QY	42	-----TSREGGGGPGPEASLNKGLAKHLLSGLDRLCLRLKREALAWAQREG-Q	91	
DB	120	ITLGGPTEBTEENTYTGFEVS--DPMSTVIEELGKREVTPPKYRELL--AKKEGIT	175	
QY	92	GPAMTEDSPGI PHCCSRCHHGLFNTHWRCSHRCLCVACGRAGAKKREKRGTSQEQHT	151	
DB	176	GPPADSSKPIGPDDAIDALSSDF-----TCGSPTAAGKTKKEESTEVLC	220	
QY	152	DCAQBAQHAACSL-----ILTFVSSQALAEIUSTVM--HOAWAKFDIRGHCFQVD	201	
DB	221	---AQSAGTVRSAAAPQEKRRKVEKDTMSDQALEALSASLGTRQAEPELDLRS--	275	
QY	202	-----ARYWAPDGGGQ-----KEPTKTPPTPPSCNGSDNRTK	236	
DB	276	EKAKEEKEKCEGDEDTIPSEYRLKPAATDKGKPLLPPEEKPKPRSELIJESDF	335	
QY	237	DIKE--ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPA--	292	

Db 336 DRSECKEKSPKTEKTEESKAAAPAP---VSEAVCRTS--MC-----SIQSAPPBPATL 384
QY 293 ---LPSDDRTITNILDST-----IAQVVERKIQEALGPGLRAGSLRGKGLSLPLSPV 341
Db 385 KGTVP--DDAVEALADSLGKKEADPEDGKPVMDKVKEA-----421
QY 342 RTRLSPPGALLWLOEPRPKHG-----FHLFOEHWROGQPVLSVSGIQKTLRLSLWGM 392
Db 422 -----KEEDREKLGEKEETIPDPYRLVEEVKDKGKPLLPKESKEQL-----462
QY 393 EALGTILGGQVQSILALGPPOPTNLDSTAFWEGFSGHPETRPKLDGSGVLLHRTILGDKDAS 452
Db 463 -----PPMSDFLLDALSEDFFSQPNASSLKFEDAKL-----AA 496
QY 453 RVQNLVSSLP 462
Db 497 AISEVVSQTP 506

RESULT 2

US-10-971-982-2
; Sequence 2, Application US/10971982
; Publication No. US20050244889A1
; GENERAL INFORMATION:
; APPLICANT: Rosenfeld, Michael G.
; APPLICANT: Glass, Christopher K.
; APPLICANT: Rose, David W.
; APPLICANT: Torchia, Joseph
; TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
; FILE REFERENCE: 6627-PA1021
; CURRENT APPLICATION NUMBER: US/10/971,982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US/09/445,353
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: PCT/US98/12263
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/049,452
; PRIOR FILING DATE: 1997-06-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312)..(3121)
; OTHER INFORMATION: "n" is any nucleotide
US-10-971-982-2

Query Match 2.8%; Score 102; DB 1; Length 1402;
Best Local Similarity 19.4%; Pred. No. 1.2;
Matches 123; Conservative 75; Mismatches 242; Indels 194; Gaps 27;
QY 22 QKRRSPLEKQLEEDSDSATSEGGGPGPEASLNKGLAKHLLSGIGDLRCLLRKERE 81
Db 579 QNSRDPQVKK--ESKSSGEVSETPRGP-LESKGKKLLQLLTCSSDDR-----625
QY 82 ALAWAREGOGPANTEDESGIPHC-----CSRCHCLFNTHW 118
Db 626 -----GHSLSLTPDLPNCKDSSVSTPSGVSSSTGTVSTSNVHGSLLQEKH 675
QY 119 RCSHSHRLCVACGRIAGAK-NREKTGSQEQHTDDCAQ-----EAGHAACSL 165
Db 676 R---ILHKLQNGNSPAEVAKITAEATGKTSSASCGETTRQEQLSPKKENALLRY 732
QY 166 ILTQFVSSQALAEIYSTVHQAWAKFDIRGCFQVDARVWAPGGGQKQKTEKTPPTPQ 225
Db 733 LLDRDDPSDLAK--ELQPQADSGDKLSQCSCSTNP-----SSGQEKDPKIKETNDE 784
QY 226 PSCNGDSNRTKDKEE--TPDSTESP---AEDGAGRSPLPCPSICELLASTAVKCLIGHD 280

Db 785 VS--GDLNLDAILGDLTSSDFYNNPTNGHGPACAKQOMFAGPSSSLGRSP-----832
QY 281 RIHMAFAPVTPALPSDDRTITNILDSTIAQVVERKIQEALGPGCL-----324
Db 833 -----QPQSVSRPPYNRAVS-LDSPSVSGSGPPVKNVSAFPGCLPKQPTILAGNPRMWSQ 885
QY 325 -RAGSLRKLGLSLPLSPVTRLSRPPGALLWLOEPRPKHGPHLPOEHWROGQPVLSVSGIQK 383
Db 886 ENYGAMGNPNRVNPNPT-----SSPG-----DWGLANSRASRMEPLASSPLGR 929
QY 384 TLRLSLWGMALGTL-----GGVQSLTALGPPOPTNLDSTAFWEGFSGHPETRPKLDGSG 438
Db 930 T-----GADYSATLPRPAMGGSVPTL-----PLRSNRL-----PGARPSLQQQQ 968
QY 439 VLLHRTILGDKDASRVQNLVSSLPPEYCAHQGKLNLSYPLGLTLHLEPQL-----492
Db 969 QQQQQQQQQQQQQQQQQQQQQQLQWRTCE-----IPMGVGNVPYSPAVQSNQPG 1015
QY 493 -WAAVGNSHRGHLGTYN---LCVEVSDLSILVHAEALPPHYRAKO-----FLSLDGG 544
Db 1016 SWPEGLMSMEQPGHGSQNRPLLRNSLDLLGPPSNAEQSGDE--RALLDQLHTFLSNTDA 1073
QY 545 EG-----LWSPGSGTSTVMHVFRAODA 566
Db 1074 TGLEIEDRALGIPELVNOGALESKQDVFQGEA 1107

RESULT 3
US-10-967-527A-14
; Sequence 14, Application US/10967527A
; Publication No. US20050256041A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian A.
; APPLICANT: Holloway, James L.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: Ztafr14, A Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor
; FILE REFERENCE: 03-17
; CURRENT APPLICATION NUMBER: US/10/967,527A
; CURRENT FILING DATE: 2004-10-18
; PRIOR APPLICATION NUMBER: 60/511,698
; PRIOR FILING DATE: 2003-10-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 447
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-967-527A-14

Query Match 2.7%; Score 101; DB 1; Length 447;
Best Local Similarity 23.8%; Pred. No. 0.34;
Matches 72; Conservative 31; Mismatches 119; Indels 80; Gaps 15;
QY 4 CQSCVQAA-GEVGVLT-GHSQKRRSRPLEBKQLEEDSDSATSEGGGPGPEASLNKGLA 61
Db 44 CQCPGPGGPEYLSGCGYTK-----DSDYGCVPCEKFKSG--82
QY 62 KHLISGLGDRCLRLKAREALAWAQREGOGPANTEDESGIPHCCHHGLFNTHWRCS 121
Db 83 -----GVQICR-RHKDCGFPFRAIVLTPGD-MENDAE-----CGPCLFGYYMLENRP 128
QY 122 HCSHRLCVAC-----GRIAGA-----GKNREKTSQBOHTDDCAOEAGHAACSL 165
Db 129 NIYGMVCSCLLAPPNTKECVATGATGASANFPCTSGSSTLSPPQHAHKLSGGCHLATAL 188
QY 166 IL---TQFVSSQALAEIYSTVMH-----QAWAKFDIRGCFQVDARVWAPGDDGQ-----212
Db 189 IIAWSTIFI--MAIALVLIIMFYILTKPSAPACCTSHPGKSVQAEQVSKDEEKEAPDNV 246
QY 213 -----QKEPTEK---TPPTPOPCNGSDSNRTKDIKEETPDSTESPAEDGAGRSPLPCPSLC 265
Db 247 VMFSEKDEFELTATPAKPTKSENDASSENEQLLSRSDSDERPA PDKQG-----SPELC 301

```
QY 266 EL 267
Db 302 LL 303

RESULT 4
US-10-821-234-921
; Sequence 921, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 921
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-921

Query Match 2.6%; Score 97.5; DB 1; Length 658;
Best Local Similarity 22.9%; Pred. No. 1;
Matches 86; Conservative 32; Mismatches 124; Indels 133; Gaps 22;

QY 330 LRKGLSLPLSPVTRLSPGALLMLQEPKHFHLPQEHWRGQPVLSVGTKLRLSL 389
Db 21 LEAALGLPVKKPLRGPRPGSLRLAE-----VSG----- 50

QY 390 WGMALGTLGGQ-VOSTALGPPQPTNLDSTAFWEGFHPETRPKIDEGSVLLHLRLTGD 448
Db 51 -----GGTGLRSALSVPFPQAGSSRAGSGTG-THTGSDPPMERGA----- 90

QY 449 KDAIRVQNLVSLPLPEYCAHQKUNLASYLPLGLTLHPLEQLMAAYGVNHRHGLTK 508
Db 91 -GAGRKLPTGRCPVTE-----GSTVQLIAP--WNAADVHSH--GDK 127

QY 509 N-LCVEVS---DLISLVHAEQLPPWYRAQKDFLSGLDGBGLW-----SGCSQTS 555
Db 128 DSQTCIRVSASPDRLPKKEEEAPLLFRTHLOAE---PHQGCWTVTEPAAMTPGNATP 183

QY 556 TVWHVFRQAQDAQIRRFQMCVPAGAGTL---EPGAPGSCYLDAGLRRRLREWGV--- 608
Db 184 P-----RTPVTPRLRLQL-LPLANTTLSTPNPDQASFPDP---PLRBEERALL 234

QY 609 -----SCWTLQAQGEAVLVPGA-PHQVQGLVSTISVTQHFSLSPETSALSAQ 655
Db 235 PRTHLOAELHQRCWTVTE---PAALTGNATPPRTQEVTPLELLEQL--PE----- 282

QY 656 LCHQASLP-PDHRM 669
Db 283 LVHATLSTPNPNQV 297

RESULT 5
US-11-015-546A-20
; Sequence 20, Application US/11015546A
; Publication No. US20050250126A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; APPLICANT: APPELEY, MARK

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24

; TITLE OF INVENTION: ztnf13, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-24
; CURRENT APPLICATION NUMBER: US/11/015,546A
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: 60/530,185
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial DNA sequence, polypeptide sequence of
; OTHER INFORMATION: the MBP-ztnf13 fusion
US-11-015-546A-20

Query Match 2.6%; Score 97; DB 7; Length 614;
Best Local Similarity 22.8%; Pred. No. 1.1;
Matches 79; Conservative 27; Mismatches 138; Indels 102; Gaps 16;

QY 54 ASLNKGLAKHLLSG--LGDRLCELLRKEREALAWAORE-----GQGPAMT-----E 97
Db 270 ASPNKLAKBFLFNLYLLTDEGLEAVNKKPLGAVALKSYEEELAKDPRIAATWENAKQGE 329

QY 98 DSPGIPHCSCRCHHGLFNTHWRCSCSHRLCVACGRI-AGAGKNREKTSQEQTDDCAQ 156
Db 330 IMPNIPO-----MSAFW-----YAVRTAVINAASGRQTVDEALKDAQTSSSH 372

QY 157 EAGHAACSLILTFVSSQALABELSTVMHQAWAKFDIRGHCFCQVDARVWAPGGGQKQEP 216
Db 373 HHHHHANSVPLVP-----RGSREFGTAGSAMIPGGATARAEP 409

QY 217 TEKTPPTPSCNGSDNRTKDIKEETPD-----STESPAEDGAG---RSLPCLPCLCELL 268
Db 410 AR--PRAGAPRTRGSGR-----EPFGGAAATTTTTPFGGAGSARAPTPAPPAPAP 461

QY 269 ASTAVKLCIGHDIRHMAFAPVTPA-----LPSDDRTITNLDSTIAQVVERKIOEKALGPGL 324
Db 462 A-----AYAPVGAATALELPASGIGHVTT-----VLRGGADGRAAAAL 502

QY 325 RAGSGLRKGL-----SLPLSPVTRLSPPGALLMLQEPKHFHHL 365
Db 503 RGAHGHAHRLPDQRHALPESP-RQRGEAGAAASQLQAALPGGLHL 547

RESULT 6
US-10-131-826A-412
; Sequence 412, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
```


; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-9

Query Match 2.5%; Score 93.5; DB 7; Length 595;
Best Local Similarity 20.6%; Pred. No. 1.9;
Matches 119; Conservative 65; Mismatches 186; Indels 209; Gaps 35;
QY 1 VTQCQSC-----VQAAGVGVLTHGSKSR-----RSPLEKOLEBEDSSATS 43
Db 119 VNSCARCFHSCVPAGMIVKFPGTAQNTVCEPASPGVSPACASPENCK---EPSSGTI 174
QY 44 EGGGGGPGPASPKNGLAKHLLSLGLDRLCR----LIRKERALAWAQRGQGPATMETS 99
Db 175 PQ--AKPTVPSPATSSASTMPVRG-GTRLAQEAASKLTRAPDSPSSVGRPSSDPGLSPTQ 231
QY 100 P---GIPHCSCRHGGLF-NTHWRCSCSHRLCVACGRIAGAGKNREKTSQEQHTDDCA 155
Db 232 PCPEGSDCKQCEPDYLDAGRCT-----ACVSCSR-----DDL 268
QY 156 QEAGHAACSLILTFVSSQALABELSTMHQAQKFDIRGHC-----FCQVDARVMAQDGG 211
Db 269 EK---TPCAW-----NSSRTCECRPGMTCATSNATNSCARCVYPICAAET-VTKQDMA 318
QY 212 QQKEPTEKPTPTQPCNGSDNSRTKDKIETPTOSTSPAB-----DGAGRSPLPCP 262
Db 319 EKDTTTEAPPLGTQPCDN-----PTPENGAPASTPTQSILLVDSQASKTILIP 367
QY 263 S-----LCELLASTAVKLCGLGHR-----IHMFA 287
Db 368 TSAPVALSSTGKVPDAGPVFWLVILVVVVGSSAPFLLC--HRRACRKRIRKQJHLCY- 424
QY 288 PVTAPALPSDDRIITNILDIIIAQVVERKIQEALGPGLRAGSGLRKGLSL--PLSPVTRL 345
Db 425 PVOTSQPK-----LELVDSPR-----RSSTQLRSGASVTEPVAERGLM 464
QY 346 SPP-----GAL-----LWQBPRLKHG-----FHLFOEHWRRQGFVLVSGIQT 384
Db 465 SQPLMETCHSVGAAYLESPLQDASPAGGSPSPDLPEPRVSTEHTN-----NKIEK- 516
QY 385 LRLSLWGMEA---LCTLGQVSLTAL-CPPQP-----TNLSTAFWEGFSHPETPKL 434
Db 517 ----IYIMKADTVIVGTVKAELPEGRGLAGPAEPLEEELEADHTPH---YPEQETEPPL 569
QY 435 DEGS-VLLHLHRTLGDKDASRVQNLVSSSLPLPEYCAHQK 472
Db 570 GSCSDVNLSEBEGKED-----PLP--TAASGK 595

RESULT 9
US-11-184-574-4
; Sequence 4, Application US/11184574
; Publication No. US2005025541A1
; GENERAL INFORMATION:
; APPLICANT: Meng, Xiang-Jin
; APPLICANT: Haqshenas, Gholamreza
; APPLICANT: Huang, Fang-Fang
; TITLE OF INVENTION: AVIAN HEPATITIS E VIRUS, VACCINES AND METHODS OF PROTECTING
; TITLE OF INVENTION: AGAINST AVIAN HEPATITIS-SPLENOMEGALY SYNDROME AND MAMMALIAN
; TITLE OF INVENTION: HEPATITIS E
; FILE REFERENCE: AM100389
; CURRENT APPLICATION NUMBER: US/11/184,574
; CURRENT FILING DATE: 2005-07-19
; PRIOR APPLICATION NUMBER: US/10/029,840
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 483

; TYPE: PRT
; ORGANISM: Hepatitis E virus
US-11-184-574-4

Query Match 2.5%; Score 93; DB 7; Length 483;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 88; Conservative 39; Mismatches 120; Indels 140; Gaps 24;
QY 285 AFAPVTAPALPSDDRIITNILDIIIAQV-VERKIQEALGPGLRAGSGLRKGLSLPLSPVRT 343
Db 6 AAAPAGVLRAPDDTVGGLDIPPAHTDVAALVTAETAG-----HAPLE--LA 50
QY 344 RLSPPGALLWLOPRPKHGFHLFQEHWRQG---QPVLVSGIQKTLRLSLMG----- 391
Db 51 AINPPGPVL-----EQGLLYMPARLDGRDEVVKQLQSLDTHVHCLRAAP 92
QY 392 ---MEALGITLGGVQSLTALGPPQPTNLDSTA-FWE--GFSPHET-----RPKLD 435
Db 93 TSLAVINTLVGRYKATKLPVEYDLMMDTIAQFWHHIGPINFSTLEYAEMCEAMLSKGQ 152
QY 436 EGSVLLHRLHRTLGDKDASRV---QNLVSSSLPLPEYCAHQKLNLSYLPGLTLHPLBPQL 492
Db 153 DGS-LIVHLDLQADCSRITFFQDKCAKFTLDDPVAH-GKVG-----QGISAMP----- 199
QY 493 WAAVGVNSHRHGLGTKNLCVEVSDLSILVHBAQALPPWYRA-QKDFLSGLDGBGLWSPG 551
Db 200 -----KTLG-----ALFGPWFAIEKHLVAGL-----PPG 224
QY 552 SQTSTVHWVFRADAQRIIRRFLOMVC-PAGAGTLEPGAAG-----SCYLDAGLRRRL 602
Db 225 YVYG---DLTYEADLHR-----SVLCAPAGHLVFEFDSFTQNNVSLDLECELMRRF 276
QY 603 -REWGV-----SCWTLLOAPGEAV 621
Db 277 GMPDMVVALYHLVRSYV-LLVAPKEAL 302

RESULT 10
US-10-821-234-1514
; Sequence 1514, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1514
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1514

Query Match 2.5%; Score 93; DB 1; Length 915;
Best Local Similarity 21.8%; Pred. No. 3.6;
Matches 157; Conservative 61; Mismatches 257; Indels 246; Gaps 37;
QY 54 ASLNGKLAKHLISGLDRLCLRLKEREALAWAQRGQGPATMETSPIPHCCSRCHHGL 113
Db 7 AGLDEGALRKLLEVTD-----LAERRIRSAIRELQCELEEREALASKRFAERQD 60
QY 114 FNTHWRCSCSHRLC---VACGRIAGAGKNREKTSQEQHTDDCAQEAQ-----HAA 162
Db 61 NKENW--LHSQQREARQAALARLAG---QLESMDNVEELT-ALLRSAGEYBERKLTRAA 114
QY 163 CSLILTQFVSSQNALA-ELSTVMHQAWAKFDIRG---HCF--CQVDARVMAQDGGQKEP 216


```
QY 614 LOAPGBAVLPAGAHQVQGLVS-----TISVTQHPLSPETSALSAQ 655
Db 535 -----EVLGSVGP---TFLVSCDRPSTGGDEAPTPAPLLRLAQHKMKWPSVSSLDSA 595
QY 656 L-----CHQASLPPDH 667
Db 586 LESSPSLHSPAD--PSH 600

RESULT 13
US-10-984-645-2
; Sequence 2, Application US/10984645
; Publication No. US20050244386A1
; GENERAL INFORMATION:
; APPLICANT: Habener, Joel
; APPLICANT: Zulewski, Hendrik
; APPLICANT: Abraham, Elizabeth
; APPLICANT: Vallejo, Mario
; TITLE OF INVENTION: METHOD OF TRANSPLANTING IN A MAMMAL AND TREATING DIABETES MELLITU
; TITLE OF INVENTION: BY ADMINISTERING A PSEUDO-ISLET LIKE AGGREGATE DIFFERENTIATED FR
; TITLE OF INVENTION: A NESTIN-POSITIVE PANCREATIC STEM CELL
; FILE REFERENCE: 3284/1223
; CURRENT APPLICATION NUMBER: US/10/984,645
; CURRENT FILING DATE: 2004-11-09
; PRIOR APPLICATION NUMBER: US 09/731,255
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215,109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/239,880
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-645-2

Query Match 2.4%; Score 89.5; DB 1; Length 1618;
Best Local Similarity 23.5%; Pred. No. 14;
Matches 79; Conservative 23; Mismatches 121; Indels 113; Gaps 18;

QY 8 VQAAGVGVLTHGHSOKSRSPLEEKOLEEEDSSATSEEGGGGPEASLNKGLAKHLLSG 67
Db 1105 VGLGDPGHLT-----REEVNPPLEESLEAKRVQGLEGP-----RKLEE--AGG 1149
QY 68 LGDRLCLL-----RKEREALAWAQREGQGPAMT----- 96
Db 1150 LGTEFSELPGKSRDPWEPREGRESEAEAPRAEAEAPFAETLHGTSDAPSPWPLGSEE 1209
QY 97 -----EISPGI-PHCCSRCHHGLFNTHWRCSHCHSLRCVACGRIA 135
Db 1210 ABEEDVPVLVSPSTYPTILEAPGLQPOA-----EGSQEASWGVQ-----GR-A 1253
QY 136 GAGK---NREKTSQSQHDDCAQEAAGHAACSLLTQFVSSQALAE-----LST 181
Db 1254 EAKVSEBQELGSGE--IPGLQEEGES-----RESEEDLGETLPDSTPLGYLRS 1306
QY 182 VMHQAWAKFDIRGHCFQVDARVWAPGDGQQ--KEPTEKTP-PTQPSCNGDSNRKTKDI 238
Db 1307 PTPSPRTPLESRGHPKLTGKCGMDPAVLASEGLEPSEPERGESEEGEECGGRDS-----DL 1362
QY 239 KEETPD-STESPAEDGA-GRSPLPCPSLCELLASTA 272
Db 1363 SEEFEDLGTEAPPLPGVGPVEAPPLGQVPQLLDPA 1398

RESULT 14
US-11-182-946-6
; Sequence 6, Application US/11182946
; Publication No. US20050255100A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/11/182,946
; CURRENT FILING DATE: 2005-07-18
; PRIOR APPLICATION NUMBER: US/10/186,643
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: US/09/573,986
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-182-946-6

Query Match 2.4%; Score 89; DB 7; Length 415;
Best Local Similarity 19.9%; Pred. No. 2.7;
Matches 76; Conservative 51; Mismatches 171; Indels 84; Gaps 18;

QY 106 CSRCHHGLFNTHWRCSHCHSH-RLCVACGRIAGAGKREKTGSOEQHTDDCAQEAAGHAACS 164
Db 80 CKTCPHNSNEHW--NHLSTCQLCRPCDIVLGFEEVAPCTSDRKA---ECRCQPG-MSCV 133
QY 165 LILTOPVSSQALAEISTVMHQAWAKPDIRGHCF-CQVDARVWAPGDGQQKEPTEKTPPT 223
Db 134 YLDNECVHCE---EERLVLCQPGTEAEVTDIMTDVNCVPCPKPGHFQNTSSPRARCQPH 190
QY 224 PQPSCNGDSNRKTKDIKEETPDST-----ESPAEDGA-----CRSPLPC- 261
Db 191 TRCEIQG-----LVEAAPGTSYSDTICKNPPPEPGAMLLALLLALLSLVFLFTTTLACA 243
QY 262 ----PSLCELLASTAVKLCGLGHDRHMAFAPVTPALPSPDDRITNLDLSIAQVVERKIQE 317
Db 244 WNRHPSLCRKLGLTLLKRHPGEGE-----SPCPAPRAPDPHFDLAEPLLP--MSGDLS 295
QY 318 KALGPGLRAGSLRKLGLSLPLSP-----VTRLSPFGALLMLQEPKPKHGFHLFQBHWQ 372
Db 296 SPAGP-----PTAPSLVEEVVQQSP---LVQARELEAEPEGEHGHQVAHGAN 338
QY 373 GQPLVSGIQTKTLRLSLWGMWALGTLGGQVQSLTALGPPQPTNLDSTAFWEGSHPE--T 430
Db 339 GIHTVGSVTVTGNIIYINGPVLGTRGPGDPAPPEPPYPTP-----EEGAPGPELS 392
QY 431 RPKLDEGSV--LLLHRTLGDKD 450
Db 393 TPYQEDGKAWHLAETETLGCQD 414

RESULT 15
US-10-821-234-1556
; Sequence 1556, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 1556
; LENGTH: 419
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1556

Query Match      2.4%; Score 88.5; DB 1; Length 419;
Best Local Similarity 21.1%; Pred. No. 3;
Matches 92; Conservative 45; Mismatches 172; Indels 127; Gaps 21;

QY 67 GLGDRLCRLRKEREALAWAQREGQGFAMTEDSPGIP-----HC---CSR 108
Db 32 GUGER-C-YSEAQRCDGSDCADGTD---EEDCGCPGPHPCGAAGTSGATACYLPADR 86
QY 109 CHHGLF-----NTHWRCHSHRCLVCACRIAGAKNREKTSOE-----QHTDDCAOEAGH 160
Db 87 CNYQTFCADGADERRCRHCQ-----PGNFRCRDEKCVYETWVCDQGPDCADGSD 136
QY 161 AACSLJLITOPVSSOAL-----AELSTVMHQAWAKFDIRGHCFQCQVDAR 203
Db 137 WDCSYVLPRKVIITAAVIGSLVCGLLLVIALGCTCKLYAIRTQYSIFAPLSRMEAEIVQQ 196
QY 204 VWAPGDG---GQOKEPTEKPTFPQPSGNGDSNRTKDI-----KEETPDSTESPAEDGAG 255
Db 197 QAPPSYGQLIAQGAIPPEVDFPTENPDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRG 256
QY 256 RSPLPSPCLCELLASTAVKLCGLGHDRIHMAFAPVTPALPSDDRITNILDIIIAQVVERKI 315
Db 257 R-----LMRRLVRRL-----RRWGLLPRTNTPARASEAR-----SQVTPSAA 293
QY 316 QEKAL-----GPGLRAGS-GLRKGLSLPLSPVTRL-----SPPGALLWLQEPKPKGHFLF 366
Db 294 PLEALDGGTGPAREGGAVGGQGEQAPPLFIKAPLPSASTSPATTVPPEAGPLPSLPL- 352
QY 367 QEHWROQPVLVSGIQKTLRLSLMGMEALGTGQVQSLSLALGPPQPTNLDSTAFWEGFS 426
Db 353 -----EPSLLSGVVQALR-----CRL-----LPSLGGPPGPTRSPPG-----P 384
QY 427 HPETRPKLDGSGVILL 442
Db 385 HTAVLALEDEDDVLLV 400
```

Search completed: November 25, 2005, 22:17:23
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2005, 22:00:09 ; Search time 232 Seconds
(without alignments)
2107.462 Million cell updates/sec

Title: US-10-712-629B-18

Perfect score: 3704

Sequence: 1 VTQCSCVQAAGEVGLTGH.....MDRAVFOAVKAAVGAALQEAQ 693

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3682	99.4	1182	1	HAIR MOUSE
2	3683.5	98.9	1181	2	Q4QY90_MOUSE
3	3510	94.8	1181	1	HAIR RAT
4	3122.5	84.3	1187	2	Q8WNL9_MACMU
5	3095.5	83.6	1189	1	HAIR HUMAN
6	3095.5	83.6	1189	2	Q6G830_HUMAN
7	1534	41.4	342	2	Q863E6_PIG
8	840.5	22.7	709	2	Q8H276_SHEEP
9	729	19.7	1265	2	Q9P2G7_HUMAN
10	728.5	19.7	2081	2	Q5S028_HUMAN
11	728.5	19.7	2162	2	Q8N3U0_HUMAN
12	719.5	19.4	1441	2	Q5RCM5_PONPY
13	719	19.4	2210	2	Q723E7_HUMAN
14	702	19.0	2428	2	Q692K6_MOUSE
15	700.5	18.9	1417	2	Q9NYF4_HUMAN
16	700.5	18.9	1578	2	Q9BW93_HUMAN
17	700.5	18.9	1761	2	Q71BC6_HUMAN
18	700.5	18.9	1787	2	Q9UP90_HUMAN
19	679.5	18.3	989	2	Q6P9K3_MOUSE
20	653	17.6	1331	2	Q61RB8_XENLA
21	641.5	17.3	1325	2	Q521X8_CHICK
22	636.5	17.2	1658	2	Q4T2W0_TETNG
23	634	17.1	759	2	Q9BVH6_HUMAN
24	617	16.7	1334	2	Q5HZN1_XENLA
25	610.5	16.5	793	2	Q8U5V7_MOUSE
26	608.5	16.4	744	2	Q8K2J6_MOUSE
27	608.5	16.4	749	2	Q8K2K4_MOUSE
28	608.5	16.4	1323	2	Q6PCM1_MOUSE
29	608.5	16.4	1334	2	Q6ZQ57_MOUSE
30	606.5	16.4	1321	2	Q68D47_HUMAN
31	606.5	16.4	1321	2	Q53872_HUMAN

32	606.5	16.4	1338	2	Q9Y4C1_HUMAN
33	604.5	16.3	1214	1	TSGA_RAT
34	603.5	16.3	1212	2	Q9P0K0_HUMAN
35	602.5	16.3	1325	2	Q6U0T9_HUMAN
36	598.5	16.2	1321	2	Q6N050_HUMAN
37	550.5	14.9	854	2	Q9VHC5_DROME
38	549.5	14.8	710	2	Q7QCL1_ANOGA
39	539	14.6	955	2	Q5SQZ9_HUMAN
40	493.5	13.3	492	2	Q8K2A5_MOUSE
41	490.5	13.2	444	2	Q8CC22_MOUSE
42	471.5	12.7	592	2	Q8R350_MOUSE
43	458.5	12.4	359	2	Q8CSQ9_MOUSE
44	448.5	12.1	1192	2	Q8IY08_HUMAN
45	366.5	9.9	968	2	Q6NV48_MOUSE

Q9Y4C1	homo sapien
Q63679	rattus norv
Q9P0K0	homo sapien
Q6U0T9	homo sapien
Q6N050	homo sapien
Q9VHC5	drosophila
Q7QCL1	anopheles g
Q5SQZ9	homo sapien
Q8K2A5	mus musculu
Q8CC22	mus musculu
Q8R350	mus musculu
Q8CSQ9	mus musculu
Q8IY08	homo sapien
Q6NV48	mus musculu

ALIGNMENTS

RESULT 1

ID	HAIR_MOUSE	STANDARD;	PRT;	1182 AA.
AC	Q61645; Q80Y47;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Hairless protein.			
GN	Name=Hr;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=BALB/c; TISSUE=Skin;			
RX	MEDLINE=94329587; PubMed=8052649;			
RA	Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,			
RA	Stoye J.P.;			
RT	"Structure and expression of the hairless gene of mice.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	STRAIN=C57BL/6; TISSUE=Retina;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalish D.E.,			
RA	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RL	and mouse cDNA sequences.";			
CC	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-!- FUNCTION: May act as a transcription factor that could act on to			
CC	regulate one of the phases of hair growth.			
CC	-!- SUBCELLULAR LOCATION: Nuclear.			
CC	-!- TISSUE SPECIFICITY: Expressed predominantly in brain, hair			
CC	follicles and interfollicular epidermis. No expression in dermis.			
CC	-!- DISEASE: Defects in Hr are the cause of a number of pleiotropic			
CC	effects including structural abnormalities of epithelial cells in			
CC	the hair follicles, hair loss towards the end of the first hair			

growth cycle, and the failure of subsequent hair growth cycles. Older mice carrying an hr mutation have been reported to possess altered ratios of T-cell-dependent B-cell responses. Mice homozygous for hr mutation are uniquely sensitive to UV and chemically induced skin tumors.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; Z32675; CA83587.1; -; mRNA.
DR EMBL; BC049182; AAH49182.1; -; mRNA.
DR PIR; I48378; I48378.

DR MGI; MGI:96223; Hr.
DR InterPro: IPR003347; TF Jm1C.

DR SMART; SM00558; JmjC; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.

FT	COMPBIAS	535	0	Poly-Gly.
FT	CONFLICT	401	401	P -> S (in Ref. 1).
SO	SEQUENCE	1182	AA: 127193	MW: 3APARE96C6ER1341 CRC664:

Query Match	99.4%	Score 3682;	DB 1;	Length 1182;
Best Local Similarity	99.4%	Pred. No. 2.1e-23;		
Matches 689: Conservative		1: Mismatches	23;	
			Indels	0;
				Gaps

Qy 1 VTQCSCVQAAGEVGLTGHSQKRRRSPLEEKQLEEDSSATSEEGCGGPGPPASLNKGL 60

Dp 490 VTQCSCVQAAGEVGLTGHSQKRRRSPLEEKQLEEDSSATSEEGCGGPGPPASLNKGL 549

QY	61	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTESPGIPHCCSCRHGGI.FNTHWRC	120
DB	550	AKHLLSGLDRLCRLLRKEREALAWAQREGQGPAMTESPGIPHCCSCRHGGI.FNTHWRC	609

Qy	121	SHCSHRLCVACGRIAGAGKNREKTGSOEHTDDCAQEAAGHAACSLLTLTQFVSSQALAE	180
Db	610	SHCSHRLCVACGRIAGAGKNREKTGSOEHTDDCAQEAAGHAACSLLTLTQFVSSQALAE	669

Qy	181	TVNHQAWAKFDIRGHCFQOVDARVWAPGDGGOQKEPTKPTPTPOPCNGDSNRTKD	240
Db	670	TVNHQAWAKFDIRGHCFQOVDARVWAPGDGGOQKEPTKPTPTPOPCNGDSNRTKD	729

QY	ETPDSTESPAEDCAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSSDDRIT	300
241	ETPDSTESPAEDCAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSSDDRIT	300
DB	ETPDSTESPAEDCAGRSPLPCPSICELLASTAVKLCIGHDRIHMAFAPVTPALPSSDDRIT	789

Qy	301	NILDSIIAQWVERKIOEALGPGIRAGSLRGKGLSLPSVVRTLSPPCALLWIOEPRPK	360
Db	790	NILDSIIAQWVERKIOEALGPGIRAGSLRGKGLSLPSVVRTLSPPCALLWIOEPRPK	849

Qy	361	HGPHLFOEHWRCQPVLVSGIQKTLRLSLWGMEALGTLGGQVQSLTALGPPQPPTNLDSFA	420
Db	850	HGPHLFOEHWRCQPVLVSGIQKTLRLSLWGMEALGTLGGQVQSLTALGPPQPPTNLDSFA	909

Qy	421	FWGFSHPETRP	KLDEGSVLLHRT	LGDKDASRVNLYSSLP	PEYCAHQGKINLAS	YLP	480
D6 <th>910</th> <th>FWGFSHPETRP</th> <th>KLDEGSVLLHRT</th> <th>LGDKDASRVNLYSSLP</th> <th>PEYCAHQGKINLAS</th> <th>YLP</th> <th>969</th>	910	FWGFSHPETRP	KLDEGSVLLHRT	LGDKDASRVNLYSSLP	PEYCAHQGKINLAS	YLP	969

Qy	481	LGITLHPLEPQLWAAAYGVNSHRHGLTGKNCVEVSDLSILVHAEAOQLPPWYRAQKDFLS	540
Db	970	LGITLHPLEPQLWAAAYGVNSHRHGLTGKNCVEVSDLSILVHAEAOQLPPWYRAQKDFLS	1029

Qy	GLDGEGLWSPG	QST	STV	VHVFRAQDA	QOR	IRF	LOW	CPAGAGT	UE	PGAPG	SCYLDAGL	UR	600
541	GLDGEGLWSPG	QST	STV	VHVFRAQDA	QOR	IRF	LOW	CPAGAGT	UE <td>PGAPG</td> <td>SCYLDAGL</td> <td>UR</td> <td>600</td>	PGAPG	SCYLDAGL	UR	600
1030	GLDGEGLWSPG	QST	STV	VHVFRAQDA	QOR	IRF	LOW	CPAGAGT	UE <td>PGAPG</td> <td>SCYLDAGL</td> <td>UR</td> <td>1089</td>	PGAPG	SCYLDAGL	UR	1089

601 RLREEWGVCWTLLQAPGSEAVLVPAGAPHQVQGLVSTTSVTQHSLSPTSALSACLCHQG 660

Db	1090	RLRREWCVCWTLLOAPGEAVLPVAPAGHQVQGLVSTISVTOHFLSPETSALSALYHQ	1149
Qy	661	ASLPDPDHRLMYAQMDDRAVFOAVKAAVGALEAK	693
Db	1150	ASLPDPDHRLMYAQMDDRAVFOAVKAAVGALEAK	1182

RESULT 2

Q4QY90_MOUSE	
ID	Q4QY90_MOUSE PRELIMINARY; PRT; 1181 AA.
AC	Q4QY90;
DT	13-SEP-2005 (TrEMBLrel. 31, Created)
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE	Hairless protein.
GN	Name=Hr; Synonyms=hr;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC	Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	NUCLEOTIDE SEQUENCE..
RP	STRAIN=KM; TISSUE=Skin;
RC	PubMed=15955095; DOI=10.1111/j.0022-202X.2005.23744.x;
RX	Zhang J.-T., Fang S.G., Wang C.Y.;
RA	"A Novel Nonsense Mutation and Polymorphisms in the Mouse Hairless
RT	Gene.";
RT	J. Invest. Dermatol. 124:1200-1205(2005).
RL	EMBL; AY547391; AAT45233.1; -; mRNA.
DR	MG1; MG1:96223; Hr.
DR	SEQUENCE 1181 AA: 127068 MW: 53DEF4A87D8207BE CRC64:

Query Match	Score	DB 2;	Length
Best Local Similarity	98.9%	3663.5;	1181;
Matches	99.1%	Pred. No. 3.6e-234;	
Matches	587.	2: Mismatches	3: Indels
Conservative			1: Gaps

Qy 1 VTQCSCVOAAGEVGLTGHSSQKSRRLPLEEKOLEEEDSSATSEEGGGGPGPEASLNKGL 60

Dh 490 VTQCSCVOAAGEVGLTGHTOKSRRLPLEEKOLEEEDSSATS-EGGGGPGPEASLNKGL 548

	Qy	61	AHLLSGLDRLCRLLRREALAWAQREGOGPAMTESPGIPHCCSRCHHGLENTHNRC	120
	Dd	549	AHILLSGLDRLCRLLRREALAWAQREGOGPAMTESPGIPHCCSRCHHGLENTHNRC	608

Qy	121	SHCSHRLCVACGRIAGAGKNREKTSQEOHTDDCAQEAAGHAACSLILITQFVSSOALAEILS	180
Db	609	SHCSHRLCVACGRIAGAGKNREKTSQEOHTDDCAQEAAGHAACSLILITQFVSSOALAEILS	668

Qy	181	TVNHQAWAKFDIRGHCF	COVDARVWAPDGGGOKEPT	TKPTPTPTQPS	CNGDSNRTKD	IK	240
Db	669	TVNHQAWAKFDIRGHCF	COVDARVWAPDGGGOKEPT	TKPTPTPTQPS	CNGDSNRTKD	IK	728

Qy	241	ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRRIT	300
Db	729	ETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRRIT	788

Qy	301	NILDSIIAOWVERKIOEKALGPGURAGSGLRKGLSLPLSPVTRTFLSPPCALLWIQEPRPK	360
Db	789	NILDSIIAOWVERKIOEKALGPGURAGSGLRKGLSLPLSPVTRTFLSPPCALLWIQEPRPK	848

Qy	361	GGPHLFOEHWROGQPVLSVGIQKTLRLSLWGMEALGTLGGQVQSILALGPPQPPNL	DSTA	420
Db	849	GGPHLFOEHWROGQPVLSVGIQKTLRLSLWGMEALGTLGGQVQVITLALGPPQPPNL	DSTA	908

Qy	421	FWEGFSHPETRPKLD	EGSVLLHRTLGDKDASRVQNV	SSLPPEYCAHOGKLN	ASYLP	480
Db	909	FWEGFSHPETRPKLD	EGSVLLHRTLGDKDASRVQNV	SSLPPEYCAHOGKLN	ASYLP	968

Qy	481	969	Db
	LGLTLHPLEPOLWAAAYGVNSHRHGLGTGNLCVEVSDLSILVHAEAAQLPPWYRAQKDFLS	LGLTLHPLEPOLWAAAYGVNSHRHGLGTGNLCVEVSDLSILVHAEAAQLPPWYRAQKDFLS	LGLTLHPLEPOLWAAAYGVNSHRHGLGTGNLCVEVSDLSILVHAEAAQLPPWYRAQKDFLS
	540	102	102

QY 541 GLDGEGLWSPGSGTSTVMHVFRQAQDAQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAGLRR 600
 DB 1029 GLDGEGLWSPGSGTSTVMHVFRQAQDAQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAGLRR 1088
 QY 601 RLREENGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQAQLCHOG 660
 DB 1089 RLREENGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQAQLCHOG 1148
 QY 661 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 693
 DB 1149 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 1181

RESULT 3

HAIR_RAT
 ID HAIR_RAT STANDARD; PRT; 1181 AA.
 AC P97609;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hairless protein.
 GN Name=Hr;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroides; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=97141510; PubMed=8987811;
 RA Thompson C.C.;
 RT "Thyroid hormone-responsive genes in developing cerebellum include a
 novel synaptotagmin and a hairless homolog.";
 RL J. Neurosci. 16:7832-7840(1996).
 CC -!- FUNCTION: May act as a transcription factor that could act on to
 regulate one of the phases of hair growth.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 DR EMBL; U71293; AAC53018.1; ALT INIT; mRNA.
 DR Ensembl; ENSRNOG0000011427; Rattus norvegicus.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW DNA-binding; Metal-binding; Nuclear protein; Transcription;
 KW Transcription regulation; Zinc; Zinc-finger.
 FT ZN_FING 594 619 C6-type.
 SQ SEQUENCE 1181 AA; 127307 MW; 834B7029CF8E88F0 CRC64;

Query Match 94.8%; Score 3510; DB 1; Length 1181;
 Best Local Similarity 95.1%; Pred. No. 5.5e-224;
 Matches 659; Conservative 11; Mismatches 21; Indels 2; Gaps 2;

QY 1 VTQCSQVQAAGEVGLTCHSKSRSPLEEKLEEDSSATSEBGGGPGPEASLNKGL 60
 DB 491 VTQCSQVQAAGEVGLTCHSKSRSPLEEKLEEDSSATSEBGGGPGPEASLNKGL 548
 QY 61 AKHLISGLGDRCLRLRKEREALAWAQREGQGPANTEPSGIPHCSCRHGLFNTHWR 120
 DB 549 AKHLISGLGDRCLRLRKEREALAWAQREGQGPANTEPSGIPHCSCRHGLFNTHWR 608
 QY 121 SHCSHRLCVACGRIAGAGKNREKTSQEQHTDDCAQEAHAACSLILTOFVSSQALAE 180
 DB 609 SHCSHRLCVACGRIAGAGKNREKTSQEQHTDDCAQEAHAACSLILTOFVSSQALAE 668
 QY 181 TVMHQWAKFDIRGHCFQVDARVWAPDGGQOKEPTEKTPPTPQPSGNSNRKDKIKE 240

DB 669 TVMHQWAKFDIRGHCFQVDARVWAPDGGQOKEPTEKTPPQPSGNSNRKDKIKE 728
 QY 241 ETPDSTESPAEGAGSPPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIIT 300
 DB 729 ETPDSTESPAEGAGSPPLPCPSLCELLASTAVKLCIGHDRIHMAFAPVTPALPSDDRIIT 788
 QY 301 NILDSIIAQVVERKIQEAKLGPCLRAGSGLRKGLSLPLSPVTRRLSPGPGALLWLQPRPK 360
 DB 789 NILDSIIAQVVERKIQEAKLGPCLRAGSGLRKGLSLPLSPVTRRLSPGPGALLWLQPRPK 848
 QY 361 HGFHLFQEHWRQGPVLVSGIOKTLRLSLWGEALGTLCGGQVOSLTALGPQPTNLDSTA 420
 DB 849 HGFHLFQEHWRQGPVLVSGIOKTLRLSLWGEALGTLCGGQVOTLTALGPQPTSLDSTA 908
 QY 421 FWEFGSHPETRPKLDGSGVLLHRTLTGDKDSRVQNLVSSLPLPEYCAHOGKUNLASL 480
 DB 909 FWEFGSHPETRPKLDGSGVLLHRTLTGDKDSRVQNLVSSLPLPEYCAHOGKUNLASL 968
 QY 481 LGTLHLPLEPQWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPLPWYRAQKDFLS 540
 DB 969 LGTLHLPLEPQWAAVGVNSHRHGLGTKNLCVEVSDLSILVHAEALQPLPWYRAQKDFLS 1028
 QY 541 GLDGEGLWSPGSGTSTVMHVFRQAQDAQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAGLRR 600
 DB 1029 GLDGEGLWSPGSGTSTVMHVFRQAQDAQRIIRRFLOWMCPAGAGTLEPGAGSCYLDAGLRR 1088
 QY 601 RLREENGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQAQLCHOG 660
 DB 1089 RLREENGVCWTLLOAPGEAVLPAGAPHQVQGLVSTISVTOHFLSPETSALSQAQLCHOG 1148
 QY 661 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 693
 DB 1149 ASLPPDHRLMYAQMDRAVFAVKAAGALQAEK 1181

RESULT 4

Q8WNL9 MACMU
 ID Q8WNL9 MACMU PRELIMINARY; PRT; 1187 AA.
 AC Q8WNL9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hairless.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopitheciidae; Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Ahmad W., Ratteer M.S., Panteleyev A.A., Alta V.M., Sundberg J.P.,
 RA Christiano A.M.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF361864; AAL56245.1; -, mRNA.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 SQ SEQUENCE 1187 AA; 127134 MW; D1CFED4AAA76C77E CRC64;

Query Match 84.3%; Score 3122.5; DB 2; Length 1187;
 Best Local Similarity 83.6%; Pred. No. 2.8e-198;
 Matches 582; Conservative 43; Mismatches 68; Indels 3; Gaps 2;

QY 1 VTQCSQVQAAGEVGLTCHSKSRSPLEEKLEEDSSA-TSEBGGGPGPEASLNKG 59
 DB 492 LAQCSQAQAAGEGGGPGAGHFQQVRRSPGLGGLQOEDTAANSSSEBEGSGPDGRLSTG 551
 QY 60 LAKHLISGLGDRCLRLRKEREALAWAQREGQGPANTEPSGIPHCSCRHGLFNTHWR 119
 DB 552 LSKHLISGLGDRCLRLRKEREALAWAQREGQGPANTEPSGIPHCSCRHGLFNTHWR 611
 QY 120 SHCSHRLCVACGRIAGAGKNREKTSQEQHTDDCAQEAHAACSLILTOFVSSQALAE 179

Db 612 CPCSHRLCVACGRVAGAGRAREKAGSREQSTECTQEAHAAACSITLTFVSSQALAEI 671
QY 180 STVMHQAOKFDIRGHCFQOVDARVWAPGQGOKEPTKPTPTPQSCNGDSNRKTDIK 239
Db 672 STAHQVWVKFDIRGHCPQOVDARVWAPGQGOKESTQKPTPTPQSCNGDTHRYKSIX 731
QY 240 EETPDSTESPAEDGAGRSPLPCPSLCELLASTAVKJCLGHDRIHMAFAPVTPALPSSDRI 299
Db 732 EETPDSTETPAEDGAGAPLPCPSLCELLASTAVKJCLGHERIHMAFAPVTPALPSSDRI 791
QY 300 TNLDSIIAQVVERKIQEKGALGCLAGSLGKLSLPLSPVTRLSPPGALLWLOEPRP 359
Db 792 TNLDSIIAQVVERKIQEKGALGCLAGSLGKLSLPLSPVTRLSPPGALLWLOEPRP 851
QY 360 K--HGPHLFOEHWRGQPVLSVGTQKTLRLSLMGMEALGTLGGQVOSLTLGPPQPTNLD 417
Db 852 RPORGPHLFOEHWRGQPVLSVGIQRTLQNLGWTGALGALGQVQALSPLGPPQPTSLG 911
QY 418 STAFWEGFHPETRPKLDGSLVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 477
Db 912 STAFWEGFHPETRPKLDGSLVLLHRTLGDKDASRVQNLVSSLPPEYCAHOGKUNLAS 971
QY 478 YLPLGLTLHLEPOLWAAVGVNSHRGLGKTLNLCVEVSDLSILVHAEAOILPWYRAQKD 537
Db 972 YLPPGLALRPLEPOLWAAVGVNSHRGLGKTLNLCVEVSDLSILVHAEAOILPWYRAQKD 1031
QY 538 FLGSLGDEGLWSPGSGTSTVWHVFRQAQDQIRRFLOMVCVPAGAGTLEPGAPGSCYLDAG 597
Db 1032 FLGSLGDEGLWSPGSGTSTVWHVFRQAQDQIRRFLOMVCVPAGAGTLEPGAPGSCYLDAG 1091
QY 598 LRRLREWGVSWTLLQAPGAEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSALQC 657
Db 1092 LRRLREWGVSWTLLQAPGAEAVLPAGAPHQVQGLVSTISVTQHFLSPETSALSALQC 1151
QY 658 HQGASLPDPHRLMYAONDRAVQAVKAAVAGLQEAQ 693
Db 1152 HQGASLPDPHRLMYAONDRAVQAVKAAVAGLQEAQ 1187
RESULT 5
HAIR_HUMAN STANDARD; PRT; 1189 AA.
AC O43593; Q96H33; Q9NPE1;
DT 15-JUL-1999 (Rel. 38, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 13-SEP-2005 (Rel. 48, Last annotation update)
DR Hairless protein.
GN Name=HR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), AND VARIANT ALUNC ALA-1022.
RX MEDLINE=9445480; PubMed=9445480; DOI=10.1126/science.279.5351.720;
RA Ahmad W., ul Haque M.F., Brancolini V., Tsou H.C., Ul Haque S.,
RA Lam H., Aita V.M., Owen J., Deblaquiere M., Frank J.,
RA Cserhalmi-Friedman P.B., Leask A., McGrath J.A., Peacocke M.,
RA Ahmad M., Ott J., Cristiano A.M.;
RT "Alopecia universalis associated with a mutation in the human hairless gene."
RL Science 279:720-724 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), SEQUENCE REVISION TO 572 AND 774, AND TISSUE SPECIFICITY.
RC TISSUE=Peripheral blood leukocyte, and skin fibroblast;
RX MEDLINE=99162400; PubMed=10051399; DOI=10.1006/geno.1998.5699;
RA Ahmad W., Zlotogorski A., Panteleyev A.A., Lam H., Ahmad M.,
RA ul Haque M.F., Abdallah H.M., Dragan L., Cristiano A.M.;
RT "Genomic organization of the human hairless gene (HR) and identification of a mutation underlying congenital atrichia in an Arab Palestinian family."
RL Science 279:720-724 (1998).

Genomics 56:141-148(1999).
[3]
SEQUENCE REVISION TO 446 AND 584.
Christian A.M.;
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[4]
NUCLEOTIDE SEQUENCE (ISOFORMS 1 AND 2), VARIANT ALUNC ASP-1136, AND TISSUE SPECIFICITY.
RC TISSUE=Brain, Fetal brain, and Peripheral blood leukocyte;
RX MEDLINE=98409496; PubMed=9736769; DOI=10.1093/hmg/7.11.1671;
RA Cichon S., Anker M., Vogt I.R., Rohleder H., Putzstuck M., Hillmer A.,
RA Farooq S.A., Al-Dhafri K.S., Ahmad M., Haque S., Rietschel M.,
RA Propping P., Kruse R., Noethen M.M.;
RT "Cloning, genomic organization, alternative transcripts and mutational analysis of the gene responsible for autosomal recessive universal congenital alopecia."
RL Hum. Mol. Genet. 7:1671-1679(1998).
[5]
NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 198-1189 (ISOFORM 3).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
VARIANT GLN-620.
RX MEDLINE=98431781; PubMed=9758627;
RA Ahmad W., Irvine A.D., Lam H., Buckley C., Bingham E.A.,
RA Panteleyev A.A., Ahmad M., McGrath J.A., Cristiano A.M.;
RT "A missense mutation in the zinc-finger domain of the human hairless gene underlies congenital atrichia in a family of Irish travellers."
RL Am. J. Hum. Genet. 63:984-991(1998).
[7]
VARIANT GLN-620.
RX MEDLINE=21303039; PubMed=11410842;
RA Hillmer A.M., Kruse R., Betz R.C., Schumacher J., Heyn U.,
RA Propping P., Noethen M.M., Cichon S.;
RT "Variant 1859G-->A (Arg620Gln) of the 'hairless' gene: absence of RT association with papular atrichia or androgenic alopecia."
RL Am. J. Hum. Genet. 69:235-237(2001).
[8]
VARIANT ALUNC ASN-1012.
RX MEDLINE=22294554; PubMed=12406339;
RA DOI=10.1046/j.1523-1747.2002.00268.x;
RA Klein I., Bergman R., Indelman M., Sprecher E.;
RT "A novel missense mutation affecting the human hairless thyroid receptor interacting domain 2 causes congenital atrichia."
RL J. Invest. Dermatol. 119:920-922(2002).
CC -1- FUNCTION: May act as a transcription factor that could act on to regulate one of the phases of hair growth.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=Long;
CC IsoId=O43593-1; Sequence=Displayed;

CC Name=2; Synonyms=Short;
 CC IsoId=O43593-2; Sequence=VSP_004276;
 CC Name=3;
 CC IsoId=O43593-3; Sequence=VSP_007990, VSP_007991;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Strongest expression of isoforms 1 and 2 is
 CC seen in the small intestine, weaker expression in brain and colon,
 CC and trace expression is found in liver, pancreas, spleen, thymus,
 CC stomach, salivary gland, appendix and trachea. Isoform 1 is always
 CC the most abundant. Isoform 1 is exclusively expressed at low
 CC levels in kidney and testis and isoform 2 exclusively at high
 CC levels in the skin.
 CC -1- DISEASE: Defects in HR are the cause of alopecia universalis
 CC congenita (ALUNC) [MIM:203655]. ALUNC is a rare autosomal
 CC recessive form of hair loss characterized by hair follicles
 CC without hair.
 CC -1- DISEASE: Defects in HR are the cause of atrichia with papular
 CC lesions (APL) [MIM:209500]; also known as congenital atrichia. APL
 CC is an autosomal recessive disease characterized by papillary
 CC lesions over most of the body and almost complete absence of hair.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC DR EMBL; AF039196; AAC32258.3; -; mRNA.
 CC DR EMBL; AJ277249; CAB87577.2; -; Genomic DNA.
 CC DR EMBL; AJ277250; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ277251; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ277252; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ277253; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400825; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400826; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400827; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400828; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400829; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400830; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400831; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400832; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400833; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400834; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400835; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400836; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ400837; CAB87577.2; JOINED; Genomic DNA.
 CC DR EMBL; AJ277165; CAB86602.1; -; mRNA.
 CC DR EMBL; BC008946; -; NOT ANNOTATED CDS; mRNA.
 CC DR Ensembl; ENSG0000168453; Homo sapiens.
 CC DR HGNC; HGNC:5172; HR.
 CC DR MIM; 602302; -;
 CC DR MIM; 203655; -;
 CC DR MIM; 209500; -;
 CC DR GO; GO:0005634; C:nucleus; NAS.
 CC DR GO; GO:0003700; F:transcription factor activity; NAS.
 CC DR GO; GO:006355; P:regulation of transcription, DNA dependent; NAS.
 CC DR InterPro; IPR003347; TF_JmJc.
 CC DR Pfam; PF02373; JmJc; 1.
 CC DR SMART; SM00558; JmJc; 1.
 CC KW Alternative splicing; Disease mutation; DNA-binding; Metal-binding;
 CC Nuclear protein; Polymorphism; Transcription;
 CC KW Transcription regulation; Zinc; Zinc-finger.
 CC ZN FING 600 625
 CC FT VARSP LIC 979 994 LALRPLEQWMAAYGV -> AGPGEHSQRHSALPLP (in
 CC isoform 3).
 CC FT VARSP LIC 995 1189 /FTId=VSP_007990.
 CC FT VARSP LIC 1072 1126 Missing (in isoform 3).
 CC FT VARSP LIC 1072 1126 /FTId=VSP_007991.
 CC FT VARSP LIC 1072 1126 Missing (in isoform 2).
 CC FT VARSP LIC 620 620 /FTId=VSP_004276.
 CC FT VARSP LIC 620 620 R -> Q.
 CC FT VARSP LIC 1012 1012 /FTId=VAR_005265.
 CC FT VARSP LIC 1012 1012 D -> N (in ALUNC; affects binding to

thyroid hormone receptor).
 /FTId=VAR_016222.
 T -> A (in ALUNC).
 /FTId=VAR_005266.
 V -> D (in ALUNC).
 /FTId=VAR_005267.
 D -> G (in Ref. 4).
 S -> L (in Ref. 5).
 B4DC9C6292AA2018 CRC64;
 127553 AA; 127553 MM;
 Query Match 83.6%; Score 3095.5; DB 1; Length 1189;
 Best Local Similarity 83.2%; Pred. No. 1.8e-196;
 Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
 1 VTQCSCVCAAGRGVLTGHSOKSRSPLE-EKQLEEDSSATSEGGGGPGPEASLNKG 59
 494 LAQCSCQAAGEGGHACHSQVRRSPLEGGELUQBEDTATNSESSEGGGPGDRLSTG 553
 60 LAKHLLSGLDRLCLRLREREAALAWAQREGQGPANTDSPGPHCCSRCHHGLFNTHWR 119
 554 LAKHLLSGLDRLCLRLREREAALAWAQREGQGPANTDSPGPHCCSRCHHGLFNTHWR 613
 120 CSHCSHRLCVACGRVAGTGRAREKAGFQBSAECTQEAAGAACSLMLTQFVSSQALAE 179
 614 CPCSRLCVACGRVAGTGRAREKAGFQBSAECTQEAAGAACSLMLTQFVSSQALAE 673
 180 STVMQAWAKFDIRGHCFQVDARVWAPGDGGOKEPTKPTTPOSCNGSDNRKDIK 239
 674 STAMHQVWVKFDIRGHCPQADARVWAPGDGGOKEPTKPTTPOSCNGDTHRTKSIK 733
 240 EETPDSTESPAEDGAGRSPLPCPCCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 299
 734 EETPDSTESPAEDGAGRSPLPCPCCELLASTAVKLCGLHDIRHMAFAPVTPALPSDDRI 793
 300 TNILDSIIAQVVERKIQEALGPGLRAGSLKRLGLSLPSVTRUSPPGALLWLQEP 358
 794 TNILDSIIAQVVERKIQEALGPGLRAGSLKRLGLSLPSVTRUSPPGALLWLQEP 853
 359 -PKHGFHLFQEHWRGQPVLVSGIOKTLRLSLWGEALCTLGGOVSLTALGPPPTNLD 417
 854 CPRGPHLFQEHWRGQPVLVSGIOKTLRLSLWGEALCTLGGOVSLTALGPPPTNLD 913
 418 STAFMEGFSHPETPKLDEGSVLLHRTLGDKDASRVQNLVSLPLPEYCAHQKLNLAS 477
 914 STTFMEGFSNPELRLPKSDEGSVLLHRLALGDEDTSEVENLAASLPLPEYCALHGLN 973
 478 YLPLGLTLHPLPQLMAAYGVNSHRHGLGTLKNCVSEVSLISILVHAEAQLPWPYRAQKD 537
 974 YLPPGLALRPLPQLMAAYGVNSHRHGLGTLKNCVSEVSLISILVHAEAQLPWPYRAQKD 1033
 538 FLSGLDGGLWSPGSGTSTVMHVFRAQDAQRIRRFLOWMCPAGAGTLEPGAGSCYLDAG 597
 1034 FLSGLDGGLWSPGSGTSTVMHVFRAQDAQRIRRFLOWMCPAGAGTLEPGAGSCYLDAG 1093
 598 LRRRLREEMGVSCWTLLOAPGAVLVPAGAPHQVGLVSTISVTQHFLSPETSALSALQC 657
 1094 LRRRLREEMGVSCWTLLOAPGAVLVPAGAPHQVGLVSTISVTQHFLSPETSALSALQC 1153
 658 HQGASLPDPHRLMYAQMDBRAVFOAKAAVGAALQEAQ 693
 1154 HQGASLPDPHRLMYAQMDBRAVFOAKAAVGAALQEAQ 1189
 RESULT 6
 Q6GS30 HUMAN
 ID Q6GS30 HUMAN PRELIMINARY; PRT; 1189 AA.
 AC Q6GS30;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hairless protein, isoform a.
 GN Name=HR;
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=22198257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.M., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC067128; AA67128.1; -; mRNA.
 SQ SEQUENCE 1189 AA; 127494 MW; 67A4B95A01063387 CRC64;
 Query Match 83.6%; Score 3095.5; DB 2; Length 1189;
 Best Local Similarity 83.2%; Pred. No. 1.8e-196;
 Matches 579; Conservative 40; Mismatches 74; Indels 3; Gaps 2;
 QY 1 VTCOSVCAAGVGVLTHGHSKRSRPL-EEQLEEDSSATSEEGCGGPGPEASLNKG 59
 DB 494 LAQCQCAQAAGGEGGHACHSQVRRPLGGELOQEEEDTAINSSSEEGSGGDSRUSTG 553
 QY 60 LAKHLASGLGDLRLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHGLFNTHWR 119
 DB 554 LAKHLASGLGDLRLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHGLFNTHWR 613
 QY 120 CSHCSHRLCVACGRIAGAGNKRKTSQBOHTDDCAQBAHACSLILTOFVSSQALAE 179
 DB 614 CPRCSHRLCVACGRIAGAGNKRKTSQBOHTDDCAQBAHACSLILTOFVSSQALAE 673
 QY 180 STVMQHWAKFDLRGHCFCQVDARVWAPGGGQKQKTEPTPTPOPSCNGDSNRTKDIK 239
 DB 674 STAMHQVWVFXFDLRGHCFCQVDARVWAPGGGQKQKTEPTPTPOPSCNGDTHRTKSIK 733
 QY 240 EETPDSTESPAEDCAGRSPLPCPSLCBLASTAVKLCIGHDR IHMAFAPVTPALPSDDRI 299
 DB 734 EETPDSTESPAEDCAGRSPLPCPSLCBLASTAVKLCIGHDR IHMAFAPVTPALPSDDRI 793
 QY 300 TNILDSIIAQVVERKI QEKALGPGLRAGSGRLKGLSLPLSPVTRTLSPGALLWQEP- 358
 DB 794 TNILDSIIAQVVERKI QEKALGPGLRAGSGRLKGLSLPLSPVTRTLSPGALLWQEP 853
 QY 359 -PKHGFLFOHWTQGPVLVSGTQKTLRLSLWMEALGTGGQVOSI-TALGPPQPTNLD 417
 DB 854 CPRRGFLFOHWTQGPVLVSGTQKTLRLSLWMEALGTGGQVOSI-TALGPPQPTNLD 913
 QY 418 STARFEGFSHPETRPKLDGEGSVLLHRTLDGDKDASRVQNVLVSSLPPEYCAHQKLNLAS 477
 DB 914 STTFWEGFSHPETRPKLDGEGSVLLHRTLDGDKDASRVQNVLVSSLPPEYCAHQKLNLAS 973
 QY 478 YLPLGLTLHLEPQWAAVGNVSHRHLGTLTKNLCVSVSDLSILVHAEALPPWYRAQKD 537

DB 974 YLPPGLIARPLEPQWAAVGNVSHRHLGTLTKNLCVEVADLVSLVHADTFLPAWHPAQKD 1033
 QY 538 FLISGLDGEGLWSPGSGTSTVMHVFRAQDAQRIIRFLQWVCPAGAGTLEPGAGSGCYLDAG 597
 DB 1034 FLISGLDGEGLWSPGSGTSTVMHVFRAQDAQRIIRFLQWVCPAGAGALEPGAGSGCYLDAG 1093
 QY 598 LRRRLREEWGVCWTLLOQAPGEAVLPAGAPHQVGLVSTVSTVTOHFLSPETSALSQALC 657
 DB 1094 LRRRLREEWGVCWTLLOQAPGEAVLPAGAPHQVGLVSTVSTVTOHFLSPETSALSQALC 1153
 QY 658 HOGASLPPDHRMLYAMQDRAVFAQVKAAYGALQAEAK 693
 DB 1154 HOGSPLPPDCHLLYAMQDRAVFAQVKAAYGALQAEAK 1189
 RESULT 7
 Q863E6 PIG PRELIMINARY; PRT; 342 AA.
 ID Q863E6 PIG PRELIMINARY; PRT; 342 AA.
 AC Q863E6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hairless (Pigment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 OC Sus.
 OX NCBI_TaxID=9823;
 RN [1]_TaxID=9823;
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22755563; PubMed=12873232;
 RX DOI=10.1046/j.1365-2052.2003.01032.x;
 RA Fernandez A., Silio L., Noguera J.L., Sanchez A., Ovilo C.;
 RT "Linkage mapping of the porcine hairless gene (HR) to chromosome
 RT 14.";
 RL Anim. Genet. 34:317-318(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Fernandez A.I., Silio L., Rodriguez C., Noguera J.L., Sanchez A.,
 RA Ovilo C.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY279972; AAP33389.1; -; mRNA.
 DR Interpro: IPR003347; TF_JmJC.
 DR Pfam: PF02373; JmJC; 1.
 DR SMART: SM00558; JmJC; 1.
 FT NON_TER 1 1
 FT TER 342 342
 SQ SEQUENCE 342 AA; 37147 MW; 35F00261F07B4A42 CRC64;
 Query Match 41.4%; Score 1534; DB 2; Length 342;
 Best Local Similarity 84.8%; Pred. No. 1.5e-93;
 Matches 285; Conservative 15; Mismatches 36; Indels 0; Gaps 0;
 QY 329 GLRKGSLPLSPVTRTLSPGALLWQEPKPHGLFQSHWRQGPVLVSGTQKTLRLS 388
 DB 2 GLRKGSLPLSPVTRTLSPGALLWQEPKPHGLFQSHWRQGPVLVSGTQKTLRLS 61
 QY 389 LWMEALGTGGQVOSI-TALGPPQPTNLDSTARFEGFSHPETRPKLDGEGSVLLHRTLDG 448
 DB 62 LWETEALGALGGQVHALTPLGPPQSTSLGSTAFMEGFSRPTETPKSDEGSLVLLHRTLDG 121
 QY 449 KDSRVQNVLVSSLPPEYCAHQKLNLASVPLGLTLHLEPQWAAVGNVSHRHLGTLTK 508
 DB 122 EDASRVENLAASLPPEYCAHKGKLNLASVPLPPGRLPLVPQIWAAYGVSHPHRLGTLTK 181
 QY 509 NLCVSVSDLSILVHAEALPPWYRAQKDFLSGLDGEGLWSPGSGTSTVMHVFRAQDAQ 568
 DB 182 NLCVSVSDLSILVHAEALPPWYRAQKDFLSGLDGEGLWSPGSGTSTVMHVFRAQDAQ 241
 QY 569 IRRFLQWVCPAGAGTLEPGAGSGCYLDAGLRRRLREEWGVCWTLLOQAPGEAVLPAGAP 628
 DB 242 IRRFLQWVCPAGAGTLEPGAGSGCYLDAGLRRRLREEWGVCWTLLOQAPGEAVLPAGAP 301

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Howden P.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL713895; CAI10948.1; -; Genomic DNA.
 DR EMBL; AL590502; CAI10948.1; JOINED; Genomic DNA.
 DR EMBL; ENSG00000171988; Homo sapiens.
 DR Ensembl; ENSG00000171988; Homo sapiens.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW Receptor.
 SQ SEQUENCE 2081 AA; 231066 MW; 4DED8856C50CEDCE CRC64;
 Query Match 19.7%; Score 728.5; DB 2; Length 2081;
 Best Local Similarity 29.1%; Pred. No. 3.8e-39;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
 QY 61 AKHLLSGDLRLCRLLKREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
 DB 1343 SKYLDIIGDKFCQVTSKLTALSWVKDAK-IWKRAVRGVREMCDCACEATLNFNIHWVC 1401
 QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHT-DDCAQEAAGHAAACSLILTFVSSQALAE 179
 DB 1402 QKGFVVCLDYK----AKERKSSRDKELYAMWKVKQGHDKHLMPTQIIIGSVLTDL 1457
 QY 180 STVMHQAWAKFDIRGHCF----QVDARVWAPGDGQO-----KEPTE 218
 DB 1458 LDAMHTLREKYGKSHCHCTNKQNLQVGNFPTWNGVSVQLQVNLHNSKISLCWPSQOQ 1517
 QY 219 KTPPTPPSCNGSDNRKDKIEE---TPDSTESP----- 249
 DB 1518 LDAMHTLREKYGKSHCHCTNKQNLQVGNFPTWNGVSVQLQVNLHNSKISLCWPSQOQ 1575
 QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLHDIRHMAFAPV- 289
 DB 1576 QIKEREQDNSESPNGRTSPLVSNQNEQGSTLRDLTTAGKLRVGSTDAGIAFAPVYSM 1635
 QY 290 -TPALPSDDRITNILDIIIAQVVERKI-----QEKALPGRLRAGSGLR 331
 DB 1636 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIISAVDENNKLY 1695
 QY 332 KGLSLPLSPVTRTLSPGALLWLOEPRPKGPHLFQEHWRQGPVLVSGIQKTLRLSLWG 391
 DB 1696 S--DIPHSWICEK-----HILWLKDYKNSNWKLFKECWKQGPVAVSGVHKKNISLWK 1748
 QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFHPETRPKLDEGSVLLHRLTLDGKD 450
 DB 1749 AESISLDFGHDQADLLCKDSIIISNANVKFWDGFEVSKRQKNKSETVVLK-----KD 1804
 QY 451 -----ASRVQLNVSLSPLPEYCAHQKUNLASYPGLGLTLP-LQPWLAAAYGV 498
 DB 1805 WPSGEDEFTWMPARYEDLLKSLPLPEYCNPEGKFNLSHLP-GFFVRPDLGPRLCISAYG 1863
 QY 499 NSHRGH-LGTNLCVEVSDLSILVHAEALQPPMYRAQKDFLSGLDGEGLWS----- 549
 DB 1864 VAAKDHDIGTTLNLHIEVSDVNVILVYVGIAGKNGILSKAGILKKFEEDLDDILRKRLKD 1923
 QY 550 ----PGSQTSVHVHVFRAQQRIRRFQWVC-PAGAGTL---BFGAPGSCYLDAGLRRLR 601
 DB 1924 SSETPGA---LWIIYAGKVDKIREFLQKISKEGGLEVLPEHDPIRDQSVYVKNLQR 1979
 QY 602 LREWGVSCWTLQAPQEAULVPAGAPHQVQGLVSTISVTOHFLSPETSALSQALCHQGA 661
 DB 1980 LLEEVGVRTCTLIQFLGDIAIVLPAGALHQVQNFHSCIQVTEDFVSPHLVESFHLTQELR 2039
 QY 662 SLPP-----DHRMLYQAQMDRAVFOAKAAGVALQ 690
 DB 2040 LKKEINYYDKL----QVKNILYHAVKENVRLAK 2069

RESULT 11
 Q8N3U0 HUMAN
 ID Q8N3U0_HUMAN PRELIMINARY; PRT; 2162 AA.
 AC Q8N3U0;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein DKFZp761F0118 (Fragment).
 GN Name=DKFZp761F0118;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Amrydata;
 RA Ottenwaelder B., Obermaier B., Mewes H.W., Weil B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL831917; CAD38578.1; -; mRNA.
 DR Ensembl; ENSG00000171988; Homo sapiens.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 2162 AA; 240377 MW; BPE6F0F68A957108 CRC64;
 Query Match 19.7%; Score 728.5; DB 2; Length 2162;
 Best Local Similarity 29.1%; Pred. No. 4e-39;
 Matches 219; Conservative 118; Mismatches 267; Indels 149; Gaps 26;
 QY 61 AKHLLSGDLRLCRLLKREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
 DB 1424 SKYLDIIGDKFCQVTSKLTALSWVKDAK-IWKRAVRGVREMCDCACEATLNFNIHWVC 1482
 QY 121 SHCSHRLCVACGRIAGAKNREKTSQEQHT-DDCAQEAAGHAAACSLILTFVSSQALAE 179
 DB 1483 QKGFVVCLDYK----AKERKSSRDKELYAMWKVKQGHDKHLMPTQIIIGSVLTDL 1538
 QY 180 STVMHQAWAKFDIRGHCF----QVDARVWAPGDGQO-----KEPTE 218
 DB 1539 LDAMHTLREKYGKSHCHCTNKQNLQVGNFPTWNGVSVQLQVNLHNSKISLCWPSQOQ 1598
 QY 219 KTPPTPPSCNGSDNRKDKIEE---TPDSTESP----- 249
 DB 1599 NTPPKSEK--NGSSSPESDVGTDNKLTPPESQSPHLWADLAQKAREEKENKELTLEN 1656
 QY 250 -----AEDGAGR-SPLPC-----PSLCELLASTAVKLCGLHDIRHMAFAPV- 289
 DB 1657 QIKEREQDNSESPNGRTSPLVSNQNEQGSTLRDLTTAGKLRVGSTDAGIAFAPVYSM 1716
 QY 290 -TPALPSDDRITNILDIIIAQVVERKI-----QEKALPGRLRAGSGLR 331
 DB 1717 GAPSSKSGRTMPNILDIIASVVENKIPPSKTSKINVKPELKEEPEESIISAVDENNKLY 1776
 QY 332 KGLSLPLSPVTRTLSPGALLWLOEPRPKGPHLFQEHWRQGPVLVSGIQKTLRLSLWG 391
 DB 1777 S--DIPHSWICEK-----HILWLKDYKNSNWKLFKECWKQGPVAVSGVHKKNISLWK 1829
 QY 392 MEALG-TLGGQVQSLTALGPPQPTNLDSTAFWEGFHPETRPKLDEGSVLLHRLTLDGKD 450
 DB 1830 AESISLDFGHDQADLLCKDSIIISNANVKFWDGFEVSKRQKNKSETVVLK-----KD 1885
 QY 451 -----ASRVQLNVSLSPLPEYCAHQKUNLASYPGLGLTLP-LQPWLAAAYGV 498
 DB 1886 WPSGEDEFTWMPARYEDLLKSLPLPEYCNPEGKFNLSHLP-GFFVRPDLGPRLCISAYG 1944
 QY 499 NSHRGH-LGTNLCVEVSDLSILVHAEALQPPMYRAQKDFLSGLDGEGLWS----- 549
 DB 1945 VAAKDHDIGTTLNLHIEVSDVNVILVYVGIAGKNGILSKAGILKKFEEDLDDILRKRLKD 2004

550 QY 550 -----PGSQTSTVHVFAQAQRIIRRFLOMVC-PAGAGTL---EPGAPGCGYLDAGLRRR 601
 2005 DB 2005 SSEIPGA---LWHIYAGKVDKIREFLQKISKEQGLEVLPEHDPIRDQSWYVYKRLQR 2060
 602 QY 602 LREEMGVSWTLLQAPCEAVLVPAGAPHQVQGLVSTISVTHFLSPETSALSQALCHQGA 661
 2061 DB 2061 LLEBYGVRTCTLIQFLGDAIVLPAGALHQVQNFHSCIQVTDFVSPHVLVESPHLTQELR 2120
 662 QY 662 SLPP---DHRMLYAQMRAVFOAVKAAVGAALQ 690
 2121 DB 2121 LKKEEINYDDKL---QVKNILYHAVKEMVRALK 2150
 RESULT 12
 Q5RCM5_PONPY
 ID Q5RCM5_PONPY PRELIMINARY; PRT; 1441 AA.
 AC Q5RCM5;
 OS Q5RCM5;
 DT 01-FEB-2005 (TREMBlrel. 29, Created)
 DT 01-FEB-2005 (TREMBlrel. 29, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Hypothetical protein DKFZp469A034 (Fragment).
 GN Name=DKFZp469A034;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Pongo.
 NCBI_TaxID=9600;
 (1)
 RP NUCLEOTIDE SEQUENCE.
 RC The German cDNA Consortium;
 RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR858245; CAH90482.1; -, mRNA.
 DR InterPro; IPR003347; TF_JmJC.
 DR Pfam; PF02373; JmJC; 1.
 DR SMART; SM00558; JmJC; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 1441 AA; 160281 MW; 704993D5EACG177C CRC64;
 Query Match 19.4%; Score 719.5; DB 2; Length 1441;
 Best Local Similarity 29.4%; Pred. No. 9.6e-39;
 Matches 221; Conservative 115; Mismatches 270; Indels 145; Gaps 25;
 61 QY 61 AKHLLSGLDRLCLRLKEREALAWAQREGQGPAMTEDSPGPHCCSRCHHGLFNTHWRC 120
 703 DB 703 SKYILDIGDKFCQLVTSKXTALSWYKDKAK-IANKRAVRGVLEMDACEATLFIHNVVC 761
 121 QY 121 SHCSHRLCVACGRIAGAKNREKTSQEOHT-DDCAQEAAGHAACSLILTFVSSQALAE 179
 762 DB 762 QKGFVVCDCYK----AKERKSRDKELYAMKCKVQPHDHKHLMPQTIIIPGSLVITDL 817
 180 QY 180 STVMHQAWAKFIRGHCFCC-----QVDARVWAPGDCGQOKEPTKTP 221
 818 DB 818 LDAMHTLREKYGIKSHCHCTKNQNLQVGNFPTMNGVSVQLVNLNHSNISLCMPESQOQ 877
 222 QY 222 PTPQPS--CNQDSNRTKDIKEE---TPDSTESP----- 249
 878 DB 878 NTPQSKSEKNGSSPESDVGDTNKLTPPESQPLHLADLAQKAREEKENKELTLENQI 937
 250 QY 250 -----AEDGAGR-SPLPC-----PSICELLASTAVKLCIGHDRIHMAFAPV---T 290
 938 DB 938 KEEREQNSPNGRTSPVSNQNEOGSTLURDLTTTAGKLRVGSTADGIAFVYVSMGA 997
 291 QY 291 PALPSSDDRIITNLDLSIAQVVERKI-----OEKALGFLRAGSGLRKG 333
 998 DB 998 PSSKSGRTMNIILDDIIASVVENKI PPSKTSKINVKPELKEPSEISITAVDENNKLYS- 1056
 334 QY 334 LSLPLSPVTRTLPSPGALLWQBPRLKPHFLPQEHWRQCPVLVSGIKTURLSLWGM 393

```
Db 1688 DVGTNDKLTTPESQSPHLWLADLAQKAREEKENKELTLENQIKEEREQDNSESPNGRT 1747
QY 257 SPLPC-----PSICELLASTAVKLCGLGHDRIHMAFAPV-----TPALPSPDDRITNILDLSI 306
Db 1748 SPLVSONNEOGSLRDLLTTTAGKLRVGTSDAGIAFAPVYMGAPSKSGRTWPNILDDI 1807
QY 307 IAQVVERKI-----QEKALGFLRAGSLRGKLSLPLSPVTRRLSPPG 349
Db 1808 IASVVENKIPPSKTSKINVKPELKEPESEIIISAVDENNKLYS-DIPHSWICEK-----1860
QY 350 ALLWLQEPKPHGFHLFOEHWROGQPVLYSGIOKTLRLSLWGMEALG-TLGGQVQSILAL 408
Db 1861 HILWLKDYKNSSNNKLFKECKWQOPAVVSGVHKOMNISLWKAESISLDFGDHQADLLNC 1920
QY 409 GPPOPTNLDTAFWEGFSPHETPRPKLDGEGSVLLHRLTLGDKD-----ASRVQNL 457
Db 1921 KDSIISNANVKEFWDFEESVKQKNKSGTGVVLKL-----KDPGSGEDFTKMPARVEDL 1976
QY 458 VSSLPLPEYCAHQGKUNLASYLPLGLTLHP-LEPQLWAAVGVNSHRGH-LGTKNLCVETS 515
Db 1977 LKSLPLPEYCNPECKFNLAGHLP-GFFVRPDLGRLCSAYGVVAAKDHDIGTTNLHIEVS 2035
QY 516 DLISILVHAAQALPPWYRAQKDFLSGLDGEGLWS-----PGSQTSTVWHVFA 563
Db 2036 DVYNILVYVGIAGKNGTSLKAGILKKFEEEDLDILRLKLDSEIIPGA---LWHIYAG 2091
QY 564 QDAQTRRRFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRLRREEMGVSCWTLLOAPGE 619
Db 2092 KDVKIREFLQKTSKEQGLEVLPEHDPIRQDSWYVNNKLRLEEVGVTCTLIQLFLGD 2151
QY 620 AVILVPAGAHQVQGLVSTISVTQHFLSPE 648
Db 2152 AIVLPAGALHQVQNFHSCIQVTEDFVSPE 2180

RESULT 14
Q692K6_MOUSE PRELIMINARY; PRT; 2428 AA.
AC Q692K6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIAA1380 protein (fragment).
GN Name=JmJdic; Synonyms=mkIAA1380;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryonic intestinal tract;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
DR EMBL; AKI73162; BAD32440.1; -; mRNA.
DR MGI; MGI:1918614; JmJdic.
DR InterPro; IPR003347; TF_JmJc.
DR Pfam; PF02373; JmJc; 1.
DR SMART; SM00558; JmJc; 1.
FT NON_TER 1
SQ SEQUENCE 2428 AA; 269552 MW; D673669EDDE67779 CRC64;

Query Match 19.0%; Score 702; DB 2; Length 2428;
Best Local Similarity 28.7%; Pred. No. 2.6e-37;
Matches 215; Conservative 122; Mismatches 270; Indels 142; Gaps 28;

QY 61 AKHLLSGLDRLCLRLKEREALAWAQEGQGPANTEDSPIPCSCRHHGLFNTHWRC 120
```

```
Db 1691 SKYILDIIDGKFCQLVTSEKLTALSWYKDAK-IAWKRAVRGVREMDACBATEILFNHVWC 1749
QY 121 SHCSHRLCVACGRIAGAGKNREKTSQEOHT-DDCAQEAAGHAACSILITLTFVSSQALAE 179
Db 1750 RKCGFVACLDYK---AKERSSRDKELYAMWKCVKGQPHDKHLMPTQIIIFGSVLTDL 1805
QY 180 STVMHQAWAKFDIRGHCFC-----QV-----DARVWAPGQGG 212
Db 1806 LDAMHILREKYGIKSHCHCTNRQNLQGGNVPTMNGVSQVLQNLVHHSNKTSTSLP-ESQ 1864
QY 213 QKEPTEK-----TPPTQPSC-----NGDSNRRTKDI 238
Db 1865 QNSPQSKSQNGNSSPGSASTDSLTTPESQSPHLWLADLAQKREKEKOENKFTLREI 1924
QY 239 KES-TPDSTESP---AEDGAGRSPLPCPSLCELLASTAVKLCGLGHDRIHMAFAPV---T 290
Db 1925 KEGDQDQASDSPNGSTSPASQSQNEQSTLRDLTTTAGKLRVGTSDAGIAFAPVYMG 1984
QY 291 PALPSPDDRITNILDLSIIAQVVERKIQEKALGFLRAGSGLR-----KGLSLPLSPV-- 341
Db 1985 SSGKGGRTWPNILDDIIASVVENKIP-----PNKTSKINIKSEPNBEPKESLIPATDES 2039
QY 342 RTLSPPGA-----LLWLQEPKPHGFHLFOEHWROGQPVLYSGIOKTLRLSLWGMEAL 395
Db 2040 KSYRDIPIHFWICDQHILMKDYKNSSNNKLFKECKWQOPAVVSGVHKOMNISLWKAESI 2099
QY 396 G-TLGGQVQSLTALGPQPTNLDTAFWEGFSPHETPRPKLDGEGSVLLHRLTLGDKDA--- 451
Db 2100 SLDFGDHQADLLNCKDSISVNAVKEFWDFEESVKQKNKGGTGVVLKL---KDCPSG 2155
QY 452 -----SRVQNLVSSLPLPEYCAHQGKUNLASYLPLGLTLHP-LEPQLWAAVGVNSHR 502
Db 2156 EDFKAMPTRYEDFLRCLPLPEYCNPECKFNLAGHLP-GFFVRPDLGRLCSAYGVVAAK 2214
QY 503 GH-LGTKNLCVEVSDLSILVH-----AEQALPPWYRAQK---DFLSGLDGEGLWS 549
Db 2215 DHDIGTTLNLHIEASDVYVNVVYVGIAGKNGVLSKAGILKKFEEEDLDILRLKLDSEI 2274
QY 550 PGSQTSVWHVFRQAQARIRREFLQWVC-PAGAGTL---EPGAPGSCYLDAGLRRLRREE 605
Db 2275 PGA---LWHIYAGKVDKIREFLQKTSKEQGLEVLPEHDPIRQDSWYVNNKLRQLLEE 2330
QY 606 WGVSCWTLLOAPGEAVLVPAGAHQVQGLVSTISVTQHFLSPELTSALSALQCHQASLPP 665
Db 2331 YGVRACTLIQLGDAIVLPAGTLHQVQNFHSCQVTEDFVSPEHLVQSFLHTQELRLLE 2390
QY 666 ----DHRMLYAOMDRAVFOAKAAGALQ 690
Db 2391 EINYYDKL---QVKNILYHAVKEMVRALK 2416

RESULT 15
Q9NYF4_HUMAN PRELIMINARY; PRT; 1417 AA.
AC Q9NYF4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative zinc finger protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Bone marrow;
RA Lai F., Godley L.A., Fernald A.A., Oreilly B.J., Pamintuan L., Zhao N.,
RA Le Beau M.M.;
RT "cDNA cloning and genomic structure of three genes localized to human
RT chromosome band 5q31 encoding potential nuclear proteins.";
RT
```

Search completed: November 25, 2005, 22:10:39
Job time : 240 secs

```
RL Genomics 70:123-130(2000).
DR EMBL; AF251039; AAF63765.1; -; mRNA.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR003347; TF_JmJC.
DR Pfam; PF02373; JmJC; 1.
DR SMART; SM00558; JmJC; 1.
SQ SEQUENCE 1417 AA; 154720 MW; E9CF503D8654FD2B CRC64;

Query Match      18.9%; Score 700.5; DB 2; Length 1417;
Best Local Similarity 26.7%; Pred. No. 1.7e-37;
Matches 237; Conservative 120; Mismatches 291; Indels 239; Gaps 32;

QY 6 SCVQAAGEVGLTGHSHQKSRSPLEE-----KQLEEDSSATSE----- 44
Db 559 SCINVA-----PHLHKRECLERYRKFKQEODSTVACRFFHFRRLIFTRKGLRV 611
QY 45 EGGGGPG-----PEASLNGU-----AKHLLSGLDRLCLLRKEREALAWAORE 89
Db 612 EGPLSPQSDPDANNLWIPSSLAEGIDLSTSKYIILANVGDFCQLVNSEKEAMVMVEPH 671
QY 90 GQGPAMTEDSPGIPCHCSRHGLFNTHWRCSHRCLVACGRIAGAGKREKTGSOEQ 149
Db 672 -QXVANKRAVRGVREWCDCVETLFINHWCRKCGFGVCLDCYRLR---KSRPRSETEM 727
QY 150 HTDD-----CAQEAAGAACSLILTFQVSSQALAEISTVMHQAWAKFDIRGHCF- 201
Db 728 GDEEVFSLKCAKGQSHPELNPQTQIIPGTALYNIQDMVHAARGKWKIKANCPICISRON 787
QY 202 ARWAP-----GDGQOQKEPTEKPTPTPOPSCNGDSNRNRTK 236
Db 788 KSVLRPAVTNGMSQLPSINPSASSGNETTFSGGG-----PAPVTTPEPDHVPKADST--- 840
QY 237 DIKEETPDSTESPAEDGAGRSPL-----PCP-----SLCELLASTAVK----- 274
Db 841 DIRSEPLKTDSSASH--SNSSELKATRPCCPDTPAPSSALHWLADLATORAKEETKEAGS 898
QY 275 -----LCLG-----HDIRH----- 283
Db 899 LRSVLNKESHSPGLDSFNSTAKVSPITPKLFNSLLGLPTASNNKTEGSSRLDLLHSGPG 958
QY 284 -----MAPAPV-----TPALPSDDRIITNILDSIIAQVVERKIOEKA----- 323
Db 959 KLPQTPLDTGIPFPVPVSTSSAGVSKASLPNFDHIIASVVENKKTSDASKRACNLDT 1018
QY 324 LRAGSLRKLGLSLPLSPVTRTLP-----PGALLWLOEPRPKHGFHLFOEHWFQOGOP 375
Db 1019 QKEVKEMVWGLNV-----LDPHTSHSWLDCGRLLCLHDPSNKNKNWKIFRECWKQGP 1070
QY 376 VLVSGTQKTLRLSLMGMEALGTIGG--QVQSLTALGPPQPTNLDSTAFWEGFSPHETRPK 433
Db 1071 VLVSGVHKKLKSELWKPEAFSQEFQDQDVLVNCRCNAIISDVKVDFDWDGFEIICKRLR 1130
QY 434 LDEGSVLLH-----RTLGDKASRVQNLVSSLPLPYCAHQKGLNKLASYLPLGLTLH 486
Db 1131 SEDGQPMVLKLDWPPGEDFRDMMPTRFEDLMENLPLPEYTKRDKRGLNLASRLPSYFVRP 1190
QY 487 PLEPOLWAAYG-VNSHRHLGTNLCVEUSDLSILVHAEALPPWYRA-QKDFLSGLD- 543
Db 1191 DLGPKMYNAYGLITAEADVGTNTLHLDVSDAVNNVMVY--VGIPIGEGAHDEEVLKTIDE 1248
QY 544 -----GEGLWSPGSGTSTVMHVFRAODAQRIIRFLOWVCPAGAGTLEPGAGP-- 590
Db 1249 GDADEVTKQRIHDGKEKPA-----LWHIYAADAEKIRELLRKV---GEEQOENPPDHD 1301
QY 591 -----SCYLDAGLRRRLREEWGVCWTLTQAPGAEVLPVAGAPHQVQGLVSTISVTQHFL 645
Db 1302 PTHDQSWYLDQTLRKELYEYGVQGWAI VQFLGDVFI PAGAPHQVHNLVYSCIKVAEDFV 1361
QY 646 SPETSALSQQLCHQASLPPDH--RMLYAQMDRAVFOAVKAAVQALQ 690
Db 1362 SPEHVKHCFLTQEPFRLHLSNTHNTHEDKLQVKNIIYHAVKDAVGTLLK 1408
```

This Page Empty (10-10)